

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2002, 10:41:00 ; Search time 55.96 Seconds

(Without alignments)
1101.608 Million cell updates/sec

Title: US-09-913-414-4

Perfect score: 2807

Sequence: 1 MFRSGAKPTIFDQAVLVAR.....MNINGVISDRCGRDRCWPS 555

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

A.Geneseq_032802.*
1: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT.*
21: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT.*
22: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1 | 2807 | 100.0 | 555 | 21 | AA1982786 |
| 2 | 1195 | 42.6 | 580 | 20 | AAV55922 |
| 3 | 1193.5 | 42.5 | 583 | 20 | AAV55819 |
| 4 | 1178.5 | 42.0 | 588 | 17 | AAW05297 |
| 5 | 572.5 | 20.4 | 710 | 20 | AAW27215 |
| 6 | 569.5 | 20.3 | 707 | 18 | AAW22153 |
| 7 | 569.5 | 20.3 | 707 | 21 | AAV51407 |
| 8 | 567.5 | 20.2 | 710 | 20 | AAV27214 |
| 9 | 555.5 | 19.8 | 560 | 22 | AAU36153 |
| 10 | 552.5 | 19.7 | 711 | 18 | AAW22160 |
| 11 | 552.5 | 19.7 | 711 | 21 | AAV51413 |

| | | | | | | |
|----|-------|------|------|----|----------|--------------------|
| 12 | 549.5 | 19.6 | 723 | 22 | AAU33597 | Pseudomonas aerugi |
| 13 | 521.5 | 18.6 | 718 | 22 | AAU38443 | Salmonella typhi c |
| 14 | 457 | 16.3 | 571 | 22 | AAU39313 | Human polypeptide |
| 15 | 454 | 16.2 | 578 | 22 | AAU36908 | Staphylococcus aur |
| 16 | 453 | 16.1 | 842 | 22 | AAU09965 | Human ABC transpor |
| 17 | 450.5 | 16.0 | 574 | 21 | AAV54454 | Amino acid sequenc |
| 18 | 450 | 16.0 | 843 | 22 | AAU00011 | Human ABCB12 trans |
| 19 | 449 | 16.0 | 843 | 22 | AAU00010 | Human ABCB12 trans |
| 20 | 443 | 15.8 | 665 | 22 | ABE52752 | Escherichia coli p |
| 21 | 440 | 15.7 | 582 | 22 | AAU34522 | E. coli cellular p |
| 22 | 435 | 15.5 | 582 | 22 | AAU38413 | Salmonella typhi c |
| 23 | 422.5 | 15.1 | 548 | 22 | AAU38413 | S. epidermidis ope |
| 24 | 419.5 | 14.9 | 587 | 22 | AAU35367 | Haemophilus influe |
| 25 | 419.5 | 14.9 | 587 | 22 | AAU35367 | Haemophilus influe |
| 26 | 418.5 | 14.9 | 766 | 22 | AAU79246 | Amino acid sequenc |
| 27 | 418.5 | 14.9 | 766 | 22 | AAU79246 | Human ATP binding |
| 28 | 417.5 | 14.9 | 591 | 22 | AAU41099 | Human polypeptide |
| 29 | 415.5 | 14.8 | 766 | 22 | AAU67163 | Amino acid sequenc |
| 30 | 415.5 | 14.8 | 766 | 22 | AAU67163 | Amino acid sequenc |
| 31 | 414 | 14.7 | 1034 | 20 | AAU03792 | S. aureus polypept |
| 32 | 413.5 | 14.7 | 766 | 22 | AAU02442 | Human ATP binding |
| 33 | 409 | 14.6 | 621 | 22 | AAU92266 | C glutamicum prote |
| 34 | 409 | 14.6 | 621 | 22 | AAU92266 | Cornebacterium gl |
| 35 | 407 | 14.5 | 1247 | 22 | AAU90997 | C glutamicum prote |
| 36 | 404.5 | 14.4 | 866 | 22 | AAU71210 | Drosophila melanog |
| 37 | 404.5 | 14.4 | 603 | 22 | AAU33706 | Pseudomonas aerugi |
| 38 | 403.5 | 14.4 | 613 | 22 | AAU04887 | Micromonospora eve |
| 39 | 403.5 | 14.4 | 722 | 22 | AAU92941 | Human protein sequ |
| 40 | 403 | 14.4 | 718 | 21 | AAU71059 | Human membrane tra |
| 41 | 402 | 14.3 | 735 | 22 | AAU93546 | Human protein sequ |
| 42 | 401.5 | 14.3 | 747 | 20 | AAU89585 | Human ATP-binding |
| 43 | 401.5 | 14.3 | 747 | 21 | AAU13355 | Human ATP-binding |
| 44 | 401 | 14.3 | 717 | 19 | AAU59008 | Leucocin A ATP tra |
| 45 | 401 | 14.3 | 717 | 20 | AAU94866 | ABC transporter (1 |

ALIGNMENTS

| | |
|-----------|---|
| RESULT 1 | |
| AA1982786 | standard; protein; 555 AA. |
| XX | |
| AC | AA1982786; |
| XX | |
| DT | 22-DEC-2000 (first entry) |
| XX | |
| DE | Caulobacter crescentus ABC transporter protein, Rsad. |
| XX | |
| KW | ABC transporter protein; Rsad; ATP binding; type I secretion system; |
| KW | C-terminal secretion signal; Caulobacter host cell; biofilm formation; |
| KW | bioreactor; recombinant protein expression host; sewage treatment; |
| KW | wood pulping; fish vaccine. |
| XX | |
| OS | Caulobacter crescentus. |
| XX | |
| PN | WO200049163-A1. |
| XX | |
| PD | 24-AUG-2000. |
| XX | |
| PF | 21-FEB-2000; 2000WO-CA00173. |
| XX | |
| PR | 19-FEB-1999; 99CA-2261186. |
| XX | |
| PA | (UYBR-) UNIV BRITISH COLUMBIA. |
| XX | |
| PI | Smit J; |
| XX | |
| DR | WPI; 2000-571987/53. |
| XX | |
| PT | Caulobacter host cell for expression and secretion of a heterologous |
| PT | polypeptide, useful for treating sewage, waste water and in the pulping |
| PT | process - |

XX Claim 1; Page 35; 46pp; English.

CC The invention relates to a Caulobacter host cell for expression and
CC secretion of a heterologous polypeptide. The Caulobacter host cell
CC comprises at least one surface layer transport protein having an
CC amino acid sequence homologous to the Caulobacter crescentus Rsad or
CC Raad proteins (AAB22786, AAB22787). The host cell further comprises a
CC DNA construct comprising DNA encoding a polypeptide heterologous to a
CC surface layer (S-layer) protein of the cell, from, and operably linked
CC to, a DNA encoding a Caulobacter surface layer protein secretion signal,
CC with the proviso that when the cell comprises transport proteins having
CC the same sequence as both the Rsad and Raad proteins, the secretion
CC signal is not from the Caulobacter crescentus surface layer protein
CC (Rsad). The invention also encompasses a method for identifying a
CC Caulobacter suitable for use as a host cell for expression and secretion
CC of a heterologous polypeptide comprising the detection of homologues of
CC the rsad or raad genes using a rsad/rsad hybridisation probe. The host
CC cell is used for the expression and secretion of a heterologous
CC polypeptide. The modified Caulobacter cells may be used to treat sewage
CC and waste water. The cells may also be grown in wood pulp suspensions and
CC then used in the wood pulping process. They can also be used in fish
CC vaccines. Caulobacter species are able to form biofilms, and can attach
CC themselves to surfaces without producing the extracellular enzymes or
CC polysaccharide slimes that are characteristic of most other
CC surface-attached bacteria. Caulobacter are therefore particularly suited
CC for use in bioreactor systems. The present sequence represents the
CC Caulobacter crescentus ABC transporter protein, Rsad. This protein is
CC embedded in the inner membrane of Caulobacter crescentus and contains an
CC ATP-binding region. It recognises the C-terminal secretion signal
CC (AAB22785) of the substrate protein (e.g., Rsad), and hydrolyses ATP
CC during the transport process.

XX Sequence 555 AA;

Query Match 100.0%; Score 2807; DB 21; Length 555;
Best Local Similarity 100.0%; Pred. No. 8.9e-232;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKRSQAKPTTFDDAVICVAPAVITAMWFSFINILAVSPLYMLOVYDRVLTSMNSTL 60
DB 1 mfkrsqakpttfddavicvavpavltafwffinilavsplymlovydrvtstnvtcl 60
QY 61 IVLVICVLEFLVYGLLEALRTQVLYVRGSLKFDGVARDPFKSVLSTLSKRGIGQAFR 120
DB 61 Ivlvicvflfvygllealrtqylvrvgslkfdgvardpfksvlsltskrgigqgfr 120
QY 121 DMDQVREMTGTLAFCDAPMTFVIVYSWMLHPFFGLAIIICITITGLAVMNDNAKN 180
DB 121 dmdqvremtgltlafdcpmtfvivyswmlhpfifglaiiicititglavmndnalkn 180
QY 181 PIOMATMSIAONDAGSTLNRNAEVMKAMGWLQARWRARRDQVMOAASDAGAV 240
DB 181 pigmatmsiaondagstlnrnaevmkamgwlgarwrarrdqvwgaasdagav 240
QY 241 MSGIKVFENIVOTLLIGGAVLAIDGKISAGAMTAGSLVGRALAPIEGAVGQWKNTYGA 300
DB 241 msgikvfenvotlliggavlaidgkisagamtagslvgralapiegavgwknlyga 300
QY 301 RGAMRRLQTLMEKESADDDHNPRLPRRGVLSAEASILPRGQOQPTMQASFRIDAGAAV 360
DB 301 rgamrrlqtlmekesaddhnprrlprrgvlsaeasilprgqqpctmqasfridagaav 360
QY 361 ALVGSAGKSSILRGIVGWPCAGVIRLDGYDKOMDPKLRHVGVLPODIEILFSGT 420
DB 361 alvgsaagkssilrgivgwpcagvirldgydkomdpklrhvgvlpodieilfsgt 420
QY 421 VAQNTARFTEESQEVIEAATLAGVHEMIOSLPMGYDTAIGEGASLSGGQORLALARA 480
DB 421 vaqntarfteesqevieaatlagvhemioslpmgydtaigegasslsggqqratalara 480
QY 481 VERMALLVDEPNASLDQVGEVALMEAMKRLKAARIVITATHKVNLQAADYIMVINO 540

DB 481 vrmallvdepnasldqvgvalmeamkrlkaarivitafthkvnlqagdyimvino 540
QY 541 GVISDFGERDRCWPS 555
DB 541 gvisdfgerdrcwps 555

RESULT 2

ID AAY55922 standard; protein; 580 AA.

AC AAY55922;

DT 15-FEB-2000 (first entry)

DE Pseudomonas fluorescens ABC transporter cassette II component A.

KW ATP-binding cassette; transporter; operon; lipBCD; Serratia marcescens;

KM microbe; protein secretion.

OS Pseudomonas fluorescens.

PN JP1276172-A.

PD 12-OCT-1999.

PF 27-MAR-1998; 98JP-0080597.

PR 27-MAR-1998; 98JP-0080597.

PA (TANA) TANABE SEIYAKU CO.

DR WPI; 1999-626936/54.

DR N-PSDB; AAZ22701.

PT An ABC transporter gene - used to increase the ability of a microbe to

PS secrete proteins

XX Claim 8; Page 21-23; 28pp; Japanese.

CC This sequence represents an ATP-binding cassette (ABC) transporter
CC component A from a Pseudomonas fluorescens strain 33 and encoded by
CC the ABC transporter operon II. The operon comprises 3 genes where the
CC termination codon of the first gene overlaps with the initiation codon
CC of the second gene. The sequences of the encoded proteins have 63, 56
CC and 59% homology respectively to the lipBD proteins from Serratia
CC marcescens. The novel gene and protein can give or increase the
CC ability of a microbe for secreting a protein.

SQ Sequence 580 AA;

Query Match 42.6%; Score 1195; DB 20; Length 580;
Best Local Similarity 45.5%; Pred. No. 9.7e-94;
Matches 240; Conservative 106; Mismatches 176; Indels 6; Gaps 3;

QY 28 VFSEFINILAVSPLYMLOVYDRVLTSMNSTLIVLVICVLEFLVYGLLEALRTQVLR 87
DB 28 v fsefinilavsplymlovydrvtstnvtclvlvicvleflvygllealrtqvivr 87
QY 88 GGLKFDGVARDPFKSVLSTLSKRG---IGQAFRMDQVREMTG-GLTAFCDAPWTP 143
DB 88 gglkfdgvardpfksvlsltskrg---igqafremdqvremtg-gltafcdapwtp 143
QY 144 VFVIVSWMLHPFFGLAIIICITITGLAVMNDNAKNPIQMATMASTIAQONDAGSTLRNA 203
DB 144 vfvivswmlhpfifglaiiicititglavmndnknpiqmatmaastiaqondagstlrna 203
QY 204 EVNRAMGMWGLQARWRARRDQVMOAASDAGAVSGIKVFENIVOTLLIGGAVLA 263
DB 204 eivnrasmgwglqarwrarrdqvwgaasdagavsgikvfenvotlliggavyla 263
QY 263 eivnrasmgwglqarwrarrdqvwgaasdagavsgikvfenvotlliggavyla 263

[illegible]

| | | |
|----|-----------|--|
| XX | RESULT | 3 |
| XX | AAV55919 | |
| XX | ID | AAV55919 standard; Protein; 583 AA. |
| XX | AC | |
| XX | AAV55919; | |
| XX | DT | 15-FEB-2000 (first entry) |
| XX | DE | Pseudomonas fluorescens ABC transporter cassette component A. |
| XX | KM | ATP-binding cassette; transporter; operon; LipBD; Serratia marcescens; |
| XX | KM | microbe; protein secretion. |
| XX | OS | Pseudomonas fluorescens. |
| XX | PN | JP11276172-A. |
| XX | PN | 12-OCT-1999. |
| XX | PD | |
| XX | PF | 27-MAR-1998; 98JP-0080597. |
| XX | PR | 27-MAR-1998; 98JP-0080597. |
| XX | PA | (TANA) TANABE SEIYAKU CO. |
| XX | DR | |
| XX | DR | WPI: 1999-626936/54. |
| XX | DR | N-PSDB; AA222700. |
| XX | PT | |
| XX | PT | An ABC transporter gene - used to increase the ability of a microbe to |
| XX | PT | secrete proteins |
| XX | PS | |
| XX | PS | Claim 2; Page 14-16; 28pp; Japanese. |
| XX | CC | |
| XX | CC | This sequence represents an ATP-binding cassette (ABC) transporter |
| XX | CC | component A from an Pseudomonas fluorescens strain 33 and encoded by |
| XX | CC | an ABC transporter operon. The operon comprises 3 genes where the |
| XX | CC | termination codon of the first gene overlaps with the initiation codon |
| XX | CC | of the second gene. The sequences of the encoded proteins have 60, 44 |
| XX | CC | and 46% homology respectively to the LipBD proteins from Serratia |
| XX | CC | marcescens. The novel gene and protein can give or increase the |
| XX | CC | ability of a microbe for secreting a protein. |
| XX | Sequence | 583 AA; |
| XX | XX | |

| | | | | |
|-----------------------|-------------------|--------------------|------------|-------------|
| Query Match | 42.5%; | Score 1193.5; | DB 20; | Length 583; |
| Best Local Similarity | 44.4%; | Pred. No. 1.3e-93; | | |
| Matches 244; | Conservative 112; | Mismatches 180; | Indels 13; | Gaps 3; |

```

Oy      1 MFKSGAKPTIEDQAVLVARPAVITAMVESEFINIALVSPYMLQVYDRVLTSRNVSTL 60
        :|: | :|: | :|: | :|: | :|: | :|: | :|: | :|: | :|: | :|: | :|: |
Db      10 lfkalgdyksl-----lsgfcflalinvlmpslsymlqvydrvlssqnettl 58

```

| | | | |
|----|-----|--|-----|
| QY | 61 | IVLVITCVFLEVLVGLBALTQVLVROGKLEFDVADPFRKSVLSDTLSR-KITGQAF | 119 |
| Db | 59 | vvlltmvvgffafagvlvlevlfrstfvlsqslerrfnlrvykaafernlqrgcehgqsl | 118 |
| QY | 120 | RDMDVREEMNG-GLIARCDAPMWPVFVIVYSWMLHPFGILAIITACIIIFGLAMVNNAT | 178 |
| Db | 119 | goltlnlqrlcfpaalfafidapwfpfrylflvflfnwlvglasagavlllqiaclneylt | 178 |
| QY | 179 | KNPIOMAMASIAAONDGSTRNAEVRKAKGMNGGLOARBARBDQVMAOAAAPAGC | 238 |
| Db | 179 | kkrpjasagysqgsqclatslhmaetlcamglalrktvfkxhsqfjlqmkasttys | 238 |
| QY | 239 | AVMSGIKVFNRIVOTLLIGGAYLADOKISAGAMINGSILVGRALAPIGAVGOMKNYI | 298 |
| Db | 239 | llsstskelrlclqslvlgllvlgallvkdmctagmmiagsllmrgvlspldqlllavwqws | 298 |
| QY | 299 | GARGAMDRLQTMLEBKESADDHMLPEPRGVLSAEASILPRGAQOPTMROASFRIDAGA | 358 |
| Db | 299 | sakltagqtlndlmrefppqsgemalpparkgnvstfegysagpprrtyvpltnqvsfnlaage | 358 |
| QY | 359 | AAVALVGPASAKSSILRIRIVGWPCAGAVRILDGVYDIKKMDPEKLGHNHVVYLRDDELFs | 418 |
| Db | 359 | vlgvgaagsgskstlarlyvvgwvplclagvlvlldgadhtrwdrddl9phl9yldpddlelfs | 418 |
| QY | 419 | GTVAAONIRARFEFESEOEYEAATLAGVHEMIOSLPKCYDVALIGEGASLSGCGOROLATA | 478 |
| Db | 419 | gsiadenlarrftcdadqglvvgaaqagayvhehlrlrphdydvtlgdegsgl9gqkqyvata | 478 |
| QY | 479 | RAVFEMPALLVLDPEFNASLDVGEVALMEAMKRLKAARKRTVIFATPHVNLDAQDYIMVT | 538 |
| Db | 479 | raly9gppllvlvdepnsldtvtvgaatalasatlmqmkagstvtvltlhrsalsaqadklvl | 538 |
| QY | 539 | NOGVSTDFG 547 | |
| Db | 539 | negrlqafg 547 | |

| RESULT 4 | |
|----------|---|
| AAW05297 | |
| ID | AAW05297 standard; Protein; 588 AA. |
| XX | |
| XX | |
| AC | AAW05297; |
| XX | |
| DT | 30-DEC-1996 (first entry) |
| XX | |
| DE | Esterase secretory protein (ORF1 product). |
| XX | |
| KM | Esterase secretory gene; protein secretion. |
| XX | |
| OS | <i>Serratia marcescens</i> strain Sr41 (FERM BP-487). |

| | |
|----|-----------------------------------|
| PN | EP733707-A2. |
| XX | 25-SEP-1996. |
| PD | |
| XX | |
| PF | 20-MAR-1996; 96EP-0104429. |
| XX | |
| PR | 23-MAR-1995; 95JP-0063772. |
| XX | |
| PA | (TANA) TANABE SEIYAKU CO. |
| XX | |
| PI | Akatsuka H, Kawai E, Shibataani T |

DR WPI; 1996-444599/45.
DR N-PSDB; AAT39662.

PT New isolated esterase secretory gene from *Serratia* - used to
PT increase the productivity and extracellular secretion of esterases
XX
Claim 3; Page 15-17; 29pp; English.
PS

CC 3 Different proteins (AAW05297-99) participate in the mechanism of

PS Disclosure; Column 55-60; 96pp; English.

XX This invention describes a novel bacterial preparation (I) which
CC comprises one or more isolated and purified strains(s) of a microorganism,
CC cultured in tryptone yeast extract (TYE) broth, which produces one or
CC more RIX toxins (belonging to the family of toxins referred to as Repeats
CC in Toxins), where the strain(s) have at least one RIX toxin which is
CC cell-associated. The products of the invention have immunostimulatory,
CC antitumor, antiinflammatory, antiarthritic and antidiarrhoeal activity.
CC The bacterial preparation may be used as vaccines for the prophylaxis and
CC treatment of infectious diseases caused by strains of microorganisms
CC which produce one or more RIX toxins. The infectious diseases are swine
CC pleuropneumonia, pneumonia, septicemia, nephritis and arthritis in swine;
CC septicemia, nephritis, endocarditis and arthritis in piglets; shipping
CC fever and abortion in cattle; whooping cough, sleepy foal disease or
CC joint ill (purulent nephritis, arthritis) in foals; septicemia,
CC polyarthritis and abortion in horses; and urinary infections,
CC peritonitis, meningitis, and gastroenteritis. The bacterial preparations
CC may also be used to prepare antibodies which may be used as a means of
CC passive immunization. This sequence represents the Actinobacillus
CC pleuropneumoniae clyis protein described in the method of the invention.

XX Sequence 707 AA;

Query Match 20.3%; Score 569.5; DB 21; Length 707;

Best Local Similarity 28.4%; Pred. No. 4,5e-40;

Matches 161; Conservative 105; Mismatches 253; Indels 47; Gaps 9;

DB 10 TIFDOAVLVARPVTAMVTSFFINILALVSPMLQYDRVITSRNVSSTLIVTVICVF 69
141 twifpavikykrfiefvlvsvalqfalfaltlffqymvdklvhrfstlvdvsvall 200
QY 70 LFLVYGLLEAIRTOVLVRGGLKFDGVARDPFRKSVLDSTLS----RKGIQGAQFRMDQ 124
DB 201 vllfeivnglrltyifafstsrivdelgarifrlhlaplsyfernvgtvdvarvleq 260
QY 125 VREMTG-GLIAFCDAPTPVFTVSWMLHPFGILAIACIIIFGLAVMNDNAKTPQ 183
DB 261 lnfltgalsvldlmsffifavmwyssplklvllgslpfymgwslfslplrrld 320
QY 184 MATNASTAQNADASTLRNAEVMKAMGMLQARWRARDEQVAMQAAA----SDAGGA 239
DB 321 ekfargadngsflvesvfatnlkalavtbpmtlw----dkqlasyvsagfrytclati 376
QY 240 VMSGIKVFRNIVOTLILGGGAYLAIIDGKISAGAMIAGSILVGRALAPREGAVGOMKNYIG 299
DB 377 gqgqvqfqlkvwmvltlvgahlvsgdlsgqlafmllsgqvlaiparilaqldq 436
QY 300 ARGAMDLQOTMLRE-EKSADDMPLPEPRGVLSAEASILPPGAOQPTMRQASFR---- 353
DB 437 vglsvtrlgdvlinspleesygkialpelkgdl-----tfrnrlfrykpa 481
QY 354 -----IDAGAAVALVPSAAGKSLRLGIVGWPCAGVIRLDGYIDKMDPKL 403
DB 482 pvlindvnlsgqevvgivgrsgsktlcklqrfyipengqvlidghdialadapwl 541
QY 404 GRHVGYLPODLEFSGTVAONIARFTE-FESQEVLEATLGVHEMIOSLPMGXDTAIGE 462
DB 542 rrtgvgvqgdnvllnrsrladadagpmeklvhaaklgaghefisellegynlttve 601
QY 463 GGASLSGQRORLALARVFRPALVLDPEPNASLDQYGEVALMEAMKRLKAARTVIFA 522
DB 602 gqaglsqgqgrqfataaralvnpklllfdeatsaldysesehlmrmhqi-ckgtvtvll 660
QY 523 THKVNLLQADYIMYINGVSDSEGE 548
DB 661 ahrlstvkndrlivmekqgvlvegk 686

RESULT 8
AAAY27214
ID AAY27214 standard; Protein; 710 AA.

XX AAY27214;

XX 24-SEP-1999 (first entry)

XX Amino acid sequence of N. meningitidis protein ORF39-1.

XX Neisseria meningitidis protein; pharmaceutical; vaccine; diagnosis;
KW bacterial infection; treatment.

XX Neisseria meningitidis.

PN WO936544-A2.

PD 22-JUL-1999.

PF 14-JAN-1999; 99WO-1B00103.

PR 09-OCT-1998; 98GB-0022143.

PR 14-JAN-1998; 98GB-0000760.

PR 01-SEP-1998; 98GB-0019015.

PA (CHIR-) CHIRON SPA.

PI Grandi G, Masignani V, Piazza M, Rappuoli R, Scarlato V;

DR WPI; 1999-444400/37.

DR N-PSDB; AAX99136.

PT New protein and its nucleotide sequence, useful in vaccines or
PT diagnostic compositions for treating and/or preventing Neisseria
PS meningitidis infections

PS Claim 4; Page 75; 123pp; English.

CC The invention provides proteins (AAY27201-245) from Neisseria
CC meningitidis (strains A and B) and nucleic acid sequences (AAY99123-167)
CC encoding the proteins. Compositions comprising the protein, nucleic acid
CC or antibody specific to the protein are useful as pharmaceuticals, e.g. a
CC vaccine composition or a diagnostic composition. The composition is also
CC useful for treating or preventing an infection due to Neisseria
CC bacteria, especially Neisseria meningitidis.

SQ Sequence 710 AA;

Query Match 20.2%; Score 567.5; DB 20; Length 710;

Best Local Similarity 29.7%; Pred. No. 6,8e-40;

Matches 167; Conservative 110; Mismatches 244; Indels 41; Gaps 13;

QY 10 TIFDOAVLVARPVTAMVTSFFINILALVSPMLQYDRVITSRNVSSTLIVTVICVF 69
144 twifpavikykrfiefvlvsvalqfalfaltlffqymvdklvhrfstlvdvsvall 203
QY 70 LFLVYGLLEAIRTOVLVRGGLKFDGVARDPFRKSVLDSTLS----RKGIQGAQFRMDQ 124
DB 204 vllfeivnglrltyifafstsrivdelgarifrlhlaplsyfernvgtvdvarvleq 263
QY 125 VREMTG-GLIAFCDAPTPVFTVSWMLHP-FFGLAIACIIIFGLAV-----MN 174
DB 264 lnfltgalsvldlmsffifavmwyssclwvlasipafwsafslplrrln 323
QY 175 DNATKN-PIQMATNASTAQNADASTLRNAEVMKAMGMLQARWRARDEQVAMQAAA 233
DB 324 dkfarnadngsflvestla-----vgtvkamavepmtqtw----dnqlaayvas 369
QY 234 -----SDAGAAVMSGIKVFRNIVOTLILGGGAYLAIIDGKISAGAMIAGSILVGRALAPREG 289
DB 370 gfrvtcklavvgqgqvllqklvtvatlwlgarlvleskltyvgqlafmllsgqvaaapvir 429
QY 290 AVGOMKNY--IGARGAMDRLQOTMLR-EKSADDMPLPEPRGVLSAEASILPPGAOQPT 346
DB 430 laqlwqldqvgvisva--rlgdllnaptenasshnlalpdrlgtelrfvhrvtrvadgrlli 487

| QY | 347 | MRQSFRLDGAAGAAVLYGPSAACKSSLLKINGVWPCAAAGVIRLDGYDIKQMPDKLGRH | 406 |
|----------|--|--|-----|
| Db | 488 | Iqdlhnlrlrteagevlgvtrsgsgskstlktlvqgrllyvpeggrvlvdgndlalaapawlrtrq | 547 |
| QY | 407 | VGYLPDQIEFLFSGVVAQNINIA-RFTPEESQDEVTAAATLACVHHEMQSLPMPGYTFAIGEGA | 465 |
| Db | 548 | vgyvlgdevnllvnsrlndhnltdtgmplrlieaaklagahefnlpegyvctvvagega | 607 |
| QY | 466 | SLSGSGQRRLALAAVAFMRPALVLDEPNASLDQYGEVVAIMEAMRLKAATVIFATHK | 525 |
| Db | 608 | glsqgsgqgrlatarelatrlntprlllffeatsaldyaseralmgmqat-cantvlliahr | 666 |
| QY | 526 | VNLLAQAADYIMVINGVYISDFG | 547 |
| Db | 667 | lstvktahrlamdkgrlveag | 688 |
| RESULT 9 | | | |
| ID | AAU36153 | standard; Protein; 560 AA. | |
| AC | AAU36153: | | |
| DT | 14-FEB-2002 | (first entry) | |
| DE | Klebsiella pneumoniae cellular proliferation protein #141. | | |
| KW | antisense; prokaryotic cellular proliferation protein; | | |
| KM | antibiotic; antibacterial; drug design. | | |
| OS | Klebsiella pneumoniae. | | |
| XX | WO200170955-A2. | | |
| PN | 27-SEP-2001. | | |
| PD | | | |
| XX | | | |
| PF | 21-MAR-2001; 2001MO-US09180. | | |
| XX | | | |
| PR | 21-MAR-2000; 2000US-191078P. | | |
| PR | 23-MAY-2000; 2000US-206848P. | | |
| PR | 26-MAY-2000; 2000US-207727P. | | |
| PR | 23-OCT-2000; 2000US-242578P. | | |
| PR | 27-NOV-2000; 2000US-253625P. | | |
| PR | 22-DEC-2000; 2000US-257931P. | | |
| PR | 16-FEB-2001; 2001US-269308P. | | |
| XX | | | |
| PA | (ELIT-) ELITRA PHARM INC. | | |
| PI | Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ; | | |
| PI | Yamamoto RT, Xu HH; | | |
| DR | WPI; 2001-611495/70. | | |
| DR | N-PSDB; AAS54012. | | |
| XX | | | |
| PT | New polynucleotides for the identification and development of | | |
| PT | antibiotics, comprise sequences of antisense nucleic acids - | | |
| XX | | | |
| PS | Example 3; Seq ID No 11746; 511pp; English. | | |
| XX | | | |

| | | |
|---|--|-----|
| CC | a wide variety of organisms. The present sequence represents an | |
| CC | essential prokaryotic cellular proliferation protein. | |
| CC | Note: The sequence data for this patent did not form part | |
| CC | of the printed specification, but was obtained in electronic | |
| CC | format directly from WIPO at | |
| CC | ftp.wipo.int/pub/published_pct_sequences. | |
| XX | | |
| SQ | Sequence 560 AA: | |
| | | |
| Query Match | 19.8%; Score 555.5; DB 22; Length 560; | |
| Best Local Similarity | 30.4%; Pred. No. 5.2e-39; | |
| Matches 173; Conservative 96; Mismatches 207; Indels 93; Gaps | | 18; |
| QY | 27 MYVSEFINILAVSPLYMLQVYDRVLTSMNSTLVL----TVICVFLVYGLLEALRT 82 | |
| DB | 13 MSLASLAINLALAGLIFSMQYDRYIPQASPLLYLISGVIAAVNF---VLIKIRG 68 | |
| QY | 83 QVLVWGKLEFQGVARDPIFKSV--LDSTLSRKIGG--QAFRDMQVREPMGTGLI-AFC 137 | |
| DB | 69 hvtldllgkrgdirtdryghalrkhkthavpstsfigslreleqvremtstmtvta 128 | |
| QY | 138 DAPWPFVFIYVSMMLHPFGLIATACIIF--GLAVNNDATKRPQIMATMASTAAAND 195 | |
| DB | 129 dlpfflllqlvlvlflspwslawpaatlmlpllllgk-----knaelanxq 177 | |
| QY | 196 AGSTLRNAEV-----MKAMGMWGGLQARRDQVAVMAQAAADAG-----G 238 | |
| DB | 178 hestlnalivesigqlpdikmmgeesrlqwn-----syaltaesvqtkvtbhg 230 | |
| QY | 239 AVMSGIKVFRNIYVLTLEGAYTLADKISAGAMIASILVGRALPIE--GAVGQMK 295 | |
| DB | 231 lvsqmsl-qulllyatvavagpvlngdltlgmvaasmssrnrvapmtalcgylarwq 289 | |
| QY | 296 NTIGARGADNRQOTMLREKKSADHMPLEP---RGVLSAEASALIPGGAQPTTKRAS 352 | |
| DB | 290 qvkvakasldnlmalpvge-----gldeprlthavhgnv-----efinaef 331 | |
| QY | 353 R-----IDAGAAVALVGVSAQKSSSLRGIVGVMPACGAVIRLDGYDIK 396 | |
| DB | 332 rygkdnqpeprlikqslkagekivlgrtfgqsktlllgamgnvemtsgelrldslip 391 | |
| QY | 397 QWDPEKLRHNGVYLPDIDIEFSGVAAQNIARPTFEES--QEVLEAATLAGVHEMTOSLPMG 455 | |
| DB | 392 qddladirtnatlqgarllfhgtlrenvltgrpmasdelvkvlelsgatefynklpmg 451 | |
| QY | 456 YDTAIGEGASISGSGORQRLARAVFRMPALLVYDDEPNASIDQGVETVALMEAMKRLKAA 515 | |
| DB | 452 lshlwmeqglqsgqrgslalraralrldpnllllldpftsfidertekafeglarw-ag 510 | |
| QY | 516 KRTVIEATHTKVNLAQADIVYINGSVIS 544 | |
| DB | 511 ertlliaatkhaavlnhvdrlvlvgqla 539 | |
| RESULT 10 | | |
| AAW22160 | | |
| ID | AAW22160 standard: Protein; 711 AA. | |
| XX | | |
| AC | AAW22160; | |
| XX | | |
| DT | 16-FEB-1998 (first entry) | |
| XX | | |
| DE | ApxIITC protein. | |
| XX | | |
| KW | RTX toxin; apxICA gene; apxIBD gene; apxIIAB'C gene; apxIIABCD gene; | |
| KW | repeat in toxins toxin; cell-associated RTX toxin; vaccine production; | |
| KW | therapy; A. pleuropneumoniae infection; swine pleuropneumonia. | |
| XX | | |
| OS | Actinobacillus pleuropneumoniae. | |
| XX | | |
| PN | CA2170839-A. | |

PD 02-SEP-1996.
 XX 01-MAR-1996; 96CA-2170839.
 XX 01-MAR-1995; 95US-0396244.
 XX (UYGU-) UNIV GUELPH.
 XX MacInnes J, Mallard B, Ricciattti P, Rosendal S;
 DR WPI; 1997-245536/23.
 DR N-PSDB; AAT73220.
 XX Preparations of microorganisms producing cell-associated RTX toxins
 PT - especially for production of vaccines against swine
 PT pleuro-pneumonia
 PT
 PS Disclosure; Pages 111-113; 151pp; English.
 XX AAM22151-W22161 represent A. pleuropneumoniae RTX (repeat in toxins)
 CC toxins. These sequence are encoded by the apxICA, apxIIB, apxIIAB/C, and
 CC apxIIABD genes (see AAT73217-73220), and can be expressed by
 CC microorganisms used in the preparations of the invention. The
 CC preparations are bacterial preparations comprising one or more isolated
 CC and purified strains of a microorganism that produces one or more RTX
 CC toxins, where the strains have at least one cell-associated RTX toxin.
 CC The preparations are used for production of vaccines for the prophylaxis
 CC and treatment of infectious diseases caused by microorganisms that
 CC produce RTX toxins, where the strains have been attenuated or
 CC inactivated. The vaccines are preferably against Actinobacillus
 CC pleuropneumoniae infection (swine pleuropneumonia). It has been found
 CC that A. pleuropneumonia produces significant quantities of
 CC cell-associated RTX toxins when cultured under certain conditions, and
 CC the whole-cell protein profiles obtained from cells recovered at
 CC necropsy from the pleural fluid of infected swine. Vaccination with a
 CC bacterin prepared from heat-inactivated cultures having significant
 CC quantities of cell-associated RTX toxins give significant protection of
 CC swine against challenge with homologous strains.
 XX Sequence 711 AA;
 SO

Query Match 19.7%; Score 552.5; DB 18; Length 711;
 Best Local Similarity 28.9%; Pred. No. 1.3e-38;
 Matches 165; Conservative 117; Mismatches 238; Indels 51; Gaps 14;

QY 10 TIFDOAVLVARPAVITAMVFFINILALVSPLYMLQVYDRVLTSRNSTLVLYVICVF 69
 DB 145 twfipavikyrkifveclivsisfqlfalfpdlffgvmdkvlvnhgisltnvltalsv 204
 QY 70 LFLVYGLLEALFTOVLVNRGLKFDGVARDPFKSVSDSTLS-----RKIGGGAQFRDMQ 124
 DB 205 vvlfeivtsltylfsststriedelgaktirhllalpisfyfentrvdvtarvrelq 264
 QY 125 VREFMTG-GLIAFCDAWPTVEFIVSWMLDPFEGILAIIA--CIIFGLAVNMNATKMP 181
 DB 265 imflgqgltsvldllsfiffawmypspltlvllslpcylawsifi-----sp 317
 QY 182 IOMATYASTAON-DAGSTLRNA---EVMKAMGMGQIQAARRARDOQVAMQAAASDA 236
 DB 318 llrrrldexfarnadqsfivesvsaidtkalavrpqntlw---dkqlasysva-df 372
 QY 237 GGAVMs-----GIKVRNIVYQTLILGGGAVIAIDGKISGAMFASILLGRALAPLEGAV 291
 DB 373 rvtvlatiqgqvqlqktwmjnlwlgshlvsgdlsqglitlmisgqvlapvrila 432
 QY 292 GOMKNYIGARGAWDRLOTMLE-ERKADDMPLPEPRGYLS-----AEASILPPG 341
 DB 433 qlwqddqvgvlsitrlgvtlnspenlygklsipelfgdtaefhrrfrykpadpild-- 490
 QY 342 AOOPTMRQASFRIDAGAAVALVGPASAAGSSILLRGIVGWPACAVYIRLDGYIKQMDPE 401

DB 491 -----dnlsvkqgeivivgrsgsketlklqirfripengqvliqghdalaadpn 543
 QY 402 KLGHRVGYLPDIDELFSGTVAQNIARFTEESOE-VIEAATLAGVHEMISLPMGYDTAI 460
 DB 544 wlrrqisgvnldnvllnrstndialtdpsmsmerlyyaklagahdfiselsegyncliv 603
 QY 461 GEGGASISGGOORLALARAFAVRPALLVDEPNASLDQGEVAMRKAARRTVI 520
 DB 604 gelaqslsgqgrfialaralvmprrillldealsaldysesellmqmkl-ghrtvl 662
 QY 521 FATKVNILAQADYIVYINOGVISPGEPRD 551
 DB 663 tlhrlstcvknadrlivmekyhivegkhng 693

RESULT 11
 AAY51413
 ID AAY51413 standard; protein: 711 AA.
 AC AAY51413:
 DT 05-MAY-2000 (first entry)
 XX
 DE A. pleuropneumoniae apxIIIB protein.
 XX RTX toxin; vaccine; Repeats in Toxins; immunostimulatory; antimicrobial;
 KW antiinflammatory; antiarthritic; antiabortive; treatment; pneumonia;
 KW pleuropneumonia; septicemia; nephritis; arthritis; endocarditis;
 KW shipping fever; abortion; whooping cough; sleepy foal disease; joint ill;
 KW urinary infection; peritonitis; meningitis; gastroenteritis;
 KW passive immunization; apxIIIB.
 KW
 XX Actinobacillus pleuropneumoniae.
 OS
 XX US6013984-A.
 PN
 XX 01-FEB-2000.
 PD
 XX 23-DEC-1996; 96US-0772270.
 PF
 XX 01-MAR-1995; 95US-0396244.
 PR
 XX (UYGU-) UNIV GUELPH.
 PA
 XX Mallard B, Rosendal S, MacInnes J, Ricciattti P;
 PT WPI; 2000-146864/13.
 DR N-PSDB; AAZ88587.
 XX Bacterial preparation comprising microorganisms which produce a member
 PT of the Repeats in Toxins (RTX) family, useful for treating swine
 PT pleuropneumonia, arthritis in swine, shipping fever and abortion in
 PT cattle, and sleepy foal disease -
 PS
 PS Disclosure; Column 83-88; 96pp; English.
 XX This invention describes a novel bacterial preparation (1) which
 CC comprises one or more isolated and purified strain(s) of a microorganism,
 CC cultured in tryptone yeast extract (TYE) broth, which produces one or
 CC more RTX toxins (belonging to the family of toxins referred to as Repeats
 CC in Toxins), where the strain(s) have at least one RTX toxin which is
 CC anti-microbial. The products of the invention have immunostimulatory,
 CC antimicrobial, antiinflammatory, antiarthritic and antiabortive activity.
 CC The bacterial preparation may be used as vaccines for the prophylaxis and
 CC treatment of infectious diseases caused by strains of microorganisms
 CC which produce one or more RTX toxins. The infectious diseases are swine
 CC pleuropneumonia, pneumonia, septicemia, nephritis and arthritis in swine;
 CC septicemia, nephritis, endocarditis and arthritis in piglets; shipping
 CC fever and abortion in cattle; whooping cough, sleepy foal disease or
 CC joint ill (purulent nephritis, arthritis) in foals; septicemia,
 CC polyarthritis and abortion in horses; and urinary infections,
 CC peritonitis, meningitis, and gastroenteritis. The bacterial preparations
 CC may also be used to prepare antibodies which may be used as a means of

CC passive immunisation. This sequence represents the *Actinobacillus*
CC *pleuropneumoniae* *apxiib* protein described in the method of the
CC invention.

XX
XX
SQ Sequence 711 AA:

Query Match 19.7%; Score 552.5; DB 21; Length 711;
Best local Similarity 28.9%; Pred. No. 1.3e-38;
Matches 165; Conservative 117; Mismatches 238; Indels 51; Gaps 14;

QY 10 TIRDAVLARPAVITAMVFFINIALVSPYLMQVDRVLTNRNVLVLCVCF 69
DB 145 twfipavikyrkifvelivsisiflqfallptlffgvwmcklvhvgfscelivtalav 204
QY 70 LFLVYGLLEALRQVLYRGKLFEDGVARDPFEKSVLDSTLS-----RKGGGAFRMDQ 124
DB 205 vvlfeivlsqrlfshstsrldvelgaktlfhllalplisyfentrvgtvarvleldq 264
QY 125 VREFMTG-GLIAFCDAWTPVEVIVSMMLHPEFGILALIA--CIIIFGLVMDNATKNP 181
DB 265 lrrflggalstsvldlffiffavmypsrltvlillpoylavsifi-----sp 317
QY 182 IQMATMASTIAQN-DAGSTIRNA-----EVKAKGMNGGLQARRRAREQVAMQAAASDA 236
DB 318 lrrrldekfarnadngsfivesasldtkalavtprqmtnlw----dkqlasyvsa-df 372
QY 237 GGVVMS-----GIKVRNRYQTLILGGANLADKISACAMAGSTLVGRALAPLEGAV 291
DB 373 rvtvltlqgqgqyqlkqtwmilmvlgahlvsgdlsigqltlfmllsqvlapvrila 432
QY 292 GOMKNYIGARGAMDRLQTMLE-EKSADDMPLPEPGVLS-----AEAASILPPG 341
DB 433 qlwqddqgqvlsitrlgvdlnspenlygqklsipelfgdaafhrrfyrydpdpilld-- 490
QY 342 AQOPTMROASFRIDAGAAVALVGPSPAGSSSLRGIVGWPCAGVIRLDGYDIKQMDPE 401
DB 491 -----dvnlsykvgevgivgsvgsygtclklqlrfyfpengqvlldghdlaadpn 543
QY 402 KIGRHGYVLPDDELFGSTVPAQNIARFTEFESQE-VTEAATLIGVHEMIOSLGMGYDTAI 460
DB 544 wlrtrgvgvldqnvllnrsirtdnialtdpsmsmerlylaaklsagadfsellegyntliv 603
QY 461 GEGGASLSGGGRORLALARAVERMPALLVDEPNASLDQGEVALMEAMRLKAARKRTVI 520
DB 604 gelgaglsaggqgrlaialaralvnprrlllfdetsalsdyesehlingmqki-chgrtvi 662
QY 521 FATHKYNLLAQADYIMVINOGVISDFGERDR 551
DB 663 llahtistvknadrlivmekghivgeqkhng 693

RESULT 12
AA033597
ID AAU33597 standard; Protein; 723 AA.
XX
XX AAU33597;
XX
DT 14-FEB-2002 (first entry)
XX
DE *Pseudomonas aeruginosa* cellular proliferation protein #41.
XX
XX Antisense; prokaryotic cellular proliferation protein;
XX
XX antibiotic; antibacterial; drug design.
OS *Pseudomonas aeruginosa*.
XX
XX WO200170955-A2.
XX
XX 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US09180.
XX

PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;

DR WPI; 2001-611495/70.
N-PSDB; AAS51456.

PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -

XX
XX
XX Example 3; Seq ID No 5093; 51pp; English.

CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX
XX
SQ Sequence 723 AA;

Query Match 19.6%; Score 549.5; DB 22; Length 723;
Best local Similarity 29.3%; Pred. No. 2.4e-38;
Matches 160; Conservative 106; Mismatches 241; Indels 39; Gaps 14;

QY 22 AVITAMVFFINIALVSPYLMQVDRVLTNRNVLVLCVFLVYGLLEALR 81
DB 170 aevtae--alvanvlasasalafmqydyrvpnaafdlwlasgvalavlgvltmr 227
QY 82 TQVLVYRGKLFEDGVARDPFEKSVLDSTLSRKGIQGQAF-----RDMQVREFMTGGLIA-F 136
DB 228 ghlinvlgkrldlqstllsfyrlstrvaakpsmgfscvqefesvreflssaa11 287
QY 137 CDAPMPVPEVIVSMMLHPEFGILAIITCIITF--GLAVMMDNMTKNYIQMATMASTIAQN 194
DB 288 sdipfvafilllavlvgvhwvplvacvilm1pgl-----lqrlighsrpn 336
QY 195 DAGSTLRNAEVRAMGMMGGLQARRRAREQVAMQAAASDAGAVM-----SGIKVFR 248
DB 337 lrgamkngyllaefellevkat-rtegcclhngwellgelgelamkthtlastlsysa 395
QY 249 NIYQTLILGG-----GAYLAIDKISAGAMTAGSLVGRALPIE--GAVGOMKNYIGAR 301
DB 396 slvgqlcygvvvyfyrylsagamtvgylvacslasratlpsgsqgligrvqhtkvam 455
QY 302 GAMDRLQTMREKSSADDDHMPLEPREGVLSAEASILPPGAQOPTMROASFRIDAGAVA 361
DB 456 eglldqllmsaeqetpvgkrrfvrkerlqghyrlegvr--lahngdspvavqvalntragetva 514
QY 362 LVGPSAAGKSSSLRGIVGWPCAGVIRLDGYDIKQMDPEKGRHVGYLPDIELEFSGTV 421

```

Db      515 11gngagsgstllrl1sglllidaqg11l1ddvslctqdpadrg1gylp1qgdvalthgs1 574
QY      422 AONI-ARFTEFESQVEYIEATLAGVHEMIQSILPMGYDTAIGEGSASLSGQORRLALARA 480
Db      575 rdnl1enaal1gde1le1ctld1g1g1at1gr1h1p1g1d1m1l1-g1gn1sl1sg1q1r1g1av1lar 633
QY      481 VFEPALVLVDEPNASLDQVGEVALMEAMKRLKAARTRVFATRKVNLAAQADYIWINQ 540
Db      634 11gddp1111l1de1pta1fd1g1g1se1k1v1d1y1l1q1w1-1g1k1rt1vl1t1h1k1ks1mal1ver1av1l1q 692
QY      541 G-VISD 545
Db      693 g1v1m1d 698

RESULT 13
AAU38443
ID      AAU38443 standard; Protein; 718 AA.
XX
AC      AAU38443;
XX
DT      14-FEB-2002 (first entry)
XX
DE      Salmonella typhi cellular proliferation protein #334.
XX
KM      Antisense; prokaryotic cellular proliferation protein;
XX      antibiotic; antibacterial; drug design.
XX
OS      Salmonella typhi.
XX
PN      WO200170955-A2.
XX
PD      27-SEP-2001.
XX
PF      21-MAR-2001; 2001WO-US09180.
XX
PR      21-MAR-2000; 2000US-191078P.
XX      23-MAY-2000; 2000US-206848P.
XX      26-MAY-2000; 2000US-207727P.
XX      23-OCT-2000; 2000US-242578P.
XX      27-NOV-2000; 2000US-253625P.
XX      22-DEC-2000; 2000US-257931P.
XX      16-FEB-2001; 2001US-269308P.
XX
PA      (ELIT- ) ELITRA PHARM INC.
XX
PI      Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX      Yamamoto RT, Xu HH;
XX
DR      WPI: 2001-611495/70.
XX      N-PSDB; AASS6302.
XX
PT      New polynucleotides for the identification and development of
XX      antibiotics, comprise sequences of antisense nucleic acids -
XX
PS      Example 3; Seq ID No 14036; 511pp; English.
XX
CC      The invention relates to antisense inhibitors of genes essential to
XX      prokaryotic cellular proliferation, their use in identifying the
XX      genes, their use in the discovery of novel antibiotics, the essential
XX      genes themselves and the encoded proteins. The prokaryotes used are
XX      Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
XX      pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
XX      invention is also useful for the identification of potential new targets
XX      for antibiotic development. The antisense nucleic acids can also be used
XX      to identify proteins used in proliferation, to express these proteins,
XX      and to obtain antibodies capable of binding to the expressed proteins.
XX      The proteins can be used to screen compounds in rational drug discovery
XX      programmes. The antisense nucleic acid sequence is also useful to screen
XX      for homologous nucleic acids which are required for cell proliferation in
XX      a wide variety of organisms. The present sequence represents an
XX      essential prokaryotic cellular proliferation protein.

```

```

CC      Note: The sequence data for this patent did not form part
CC      of the printed specification, but was obtained in electronic
CC      format directly from WIPO at
CC      ftp.wipo.int/pub/published_pct_sequences.
XX
SQ      Sequence 718 AA:
XX
Query Match      18.6%; Score 521.5; DB 22; Length 718;
Best Local Similarity 28.6%; Pred. No. 5,9e-36;
Matches 164; Conservative 118; Mismatches 221; Indels 71; Gaps 19;

QY      8 KPTIEDAVLVARPAVITAVESFETINILALVSP1MYLDYIVDRVLSKRVSTIIVTVIC 67
Db      153 kpdw1relv1gdlrpy1p1w1va1f1inv1sl1ag1fsm1gyd1rv1p1ag1s1p1t1y1l1v1s1-t1g 211
QY      68 VFLEFLVYG-LLEALRTOYVLVRGKLFQGVARDIPFSVL--DSTLSRK-GIGQAFRDM 122
Db      212 vl1av1l1g1fl1r1e1ar1t1h1nd1v1l1g1k1rad1m1s1dr1v1g1h1al1l1r1ns1a1pr1st1g1s1q1r1el 271
QY      123 DQVREFMTGGLIA-FCDAFWTPVFEVIVSWMLHPFEGILATIAIIF--GLAVMNDNRK 179
Db      212 vl1av1l1g1fl1r1e1ar1t1h1nd1v1l1g1k1rad1m1s1dr1v1g1h1al1l1r1ns1a1pr1st1g1s1q1r1el 271
QY      272 eq1rem1t1s1et1s1et1v1d1p1t1f1f1m1v1l1a1l1ap1l1aw1ap1va1l1m1l1p1g1va1l1q1k1----- 326
Db      180 NP1QMATMASIAQNDAGSTLRNA-----EVKKA-----MGW-----GG 214
QY      327 ---klav1l1n1ga1a1h1e1---a1l1na1v1es1v1g1ed1k1l1m1ga1er1f1l1q1w1s1y1r1t1g1es1g 380
Db      215 LQARWRARRDEQVAMQAAADAGAVWSG1KVFERNIVOTL1LGGA1VLA1DGK1SAGAM1 274
QY      381 l1r1t1k1l1t1r1--l1s1w1-----g1m1s1v1--g1s1y1a1v1m1f1g1ap1m1v1l1eg1sm1t1g1av1 424
QY      275 AGSILVGRALAPIE---GAVGOMKNYIGARGAMDRLQTM1REK1SAD1HMP1EPR1GVL1S 331
Db      425 aa1em1l1gs1r1m1a1p1ma1n1l1o1g1va1r1aw1q1v1ka1k1m1g1nd1l1n1q1l1e1t1ch1d1s1l1h1rd1l1b1gh1y1l 484
QY      332 AEAAS1-LPPGAQOPTMRQASFRIDAGAVALVPSAAGSS1L1RG1V1G1W1PC1AA1G1V1RL 350
Db      485 f1e1n1a1q1f1r1y1h1nd1d1p1r1l1r1v1l1e1m1p1e1r1a1l1l1g1r1n1g1a1g1s1t1l1g1n1a1g1l1e1m1g1d1ar1 544
QY      391 DGVD1KQMDPEK1GRH1GY1PQD1ELFSG1VAQ1IA1FTEFESQ1E-VIEAT1AGVHEM1 449
Db      545 dn1sl1sh1md1ad1l1r1m1l1g1f1s1q1ar1l1f1g1cl1ent1l1g1p1h1and1eq1l1fd1a1l1ev1s1g1a1v1f 604
QY      450 QSLPMGYDTAIGEGSASLSGQORRLALARAVERPALVLVDEPNASLDQVGEVALMEAM 509
Db      605 r1l1a1k1y1d1h1p1m1e1g1n1l1g1s1q1r1s1l1l1a1r1m1l1r1s1p1v1l1l1d1e1p1s1a1l1d1e1h1e1---r1ef1 661
QY      510 KRLK--AAKRTVFATRKVNLAAQADYIWINOG 541
Db      662 q1l1h1q1w1l1gn1rt1vl1v1ath1rv1p1l1e1v1er1v1v1l1k1e1g 695

RESULT 14
AAU39313
ID      AAU39313 standard; Protein; 571 AA.
XX
AC      AAU39313;
XX
DT      22-OCT-2001 (first entry)
XX
DE      Human polypeptide SEQ ID NO 2458.
XX
KM      Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX      peripheral nervous system; neuropathy; central nervous system; CNS;
XX      Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX      amyotrophic lateral sclerosis; Shy-Drager Syndrome; Chemotactic;
XX      chemokine; thrombolytic; drug screening; arthritis; inflammation;
XX      leukaemia.
XX
OS      Homo sapiens.
XX
PN      WO200153312-A1.

```

XX 26-JUL-2001.
PD 26-DEC-2000: 2000MO-US34263.
XX
PF 21-JAN-2000: 2000US-0488725.
PR 25-APR-2000: 2000US-0552317.
PR 09-JUL-2000: 2000US-0598042.
PR 19-JUL-2000: 2000US-0620312.
PR 03-AUG-2000: 2000US-0653450.
PR 14-SEP-2000: 2000US-0662191.
PR 19-OCT-2000: 2000US-0693036.
PR 29-NOV-2000: 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX MPI: 2001-442253/47.
DR N-PSDB; AAI58469.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders.
PT such as central nervous system injuries -
XX
XX Example 4; SEQ ID NO 2458; 10078pp; English.
PS
XX The invention relates to human nucleic acids (AAI57798-AA161369) and
CC the encoded polypeptides (AAM38642-AA42213) with neurotropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
CC
XX
XX Sequence 571 AA:
SQ

Db 285 tnfidmemfnd---llkeetevkdldpgagplrtfkgqrliefenvhfayadgre-tlqdv 339
Qy 352 FRIDAGAAVALYVPSAAGKSSILRGIVGWPCAGVIRLDGSDYDKQMDPELGNHVGLP 411
Db 340 ftvmpgqclalvpgspgagkctllrlfrfydlssgcrltdgqdlsgvtqslshlgvvp 399
Qy 412 ODLEFSGTVAQNT--ARFEFESEVTEAATLAGVEMTOSLPMGYDPTAIGEGGASLSG 469
Db 400 gdtvllndtldadnrrygrvta-gndevaaagaagldadmafpdyrtcygvgqrlklsq 458
Qy 470 GORORLALAAVFRFPALVLDPEPNASLDVGEVALMEAMKRLKAARTVYFATHKVNL 529
Db 459 gekqvaiartllkapglilldeatsaldtsnetaqslakv-cantllvahrstlv 517
Qy 530 AQADYIMVINGVSDGERD 550
Db 518 vnadqllivkdgcylvergrhe 538
RESULT 15
ID AAU36908 standard; Protein; 578 AA.
XX
AC AAU36908;
XX
DT 14-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation protein #1078.
XX
KW Antisense; prokaryotic cellular proliferation protein;
KM antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001: 2001MO-US09180.
XX
PR 21-MAR-2000: 2000US-191078P.
PR 23-MAY-2000: 2000US-206848P.
PR 26-MAY-2000: 2000US-207727P.
PR 23-OCT-2000: 2000US-242578P.
PR 27-NOV-2000: 2000US-253625P.
PR 22-DEC-2000: 2000US-257931P.
PR 16-FEB-2001: 2001US-269308P.
XX
PA (ELIR-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR MPI: 2001-611495/70.
XX
N-PSDB; AAS54767.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 12501; 511pp; English.
PS
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery

CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX
 SQ Sequence 578 AA;

Query Match 16.2%; Score 454; DB 22; Length 578;
 Best Local Similarity 24.8%; Pred. No. 2,7e-30;
 Matches 143; Conservative 102; Mismatches 244; Indels 88; Gaps 14;

```

QY 22 AVITAMWSPFINILALVSPIMQVDRVL-----TSRWSTLIVTVICVLEFLV-- 73
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 18 attivgikigipml---pillkyaigvinnhaltdekvhltlaigiallffvlvr 74
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 74 -----YGLEALRQVLVRCGLKFDGVARDPIFKSVLDS 107
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 75 priefirgylaqwtsnkllydirkklynhlqalsarifa-----nngvqg 119
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 108 TLSRKIGGQAFRDMDOVERPMTGGLAFCDAPWTP-VFVIVSMMLHPFGILAIACII 166
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 120 vlsr-----vindveqtkdfilglml---vldcftllialsimffldvklitlaalf 169
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 167 IFGLAVNNDNATKNPIQMAT---MASIAONDAGSTLRNAEVKAMGM-----WGGL 215
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 170 lfffyiltvyffgrlklrtersqalaevgflhervgisvkvstaledneaknfdkk 229
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 216 QARWRARDEQVAAQAAADAGAVMSGKVFERNIVQTLIIGGAYLAIDGKISAGAMIA 275
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 230 ntnfltralkhttrwnays-----faaintvtdigpiivgayaiaisgslvtgltaa 282
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 276 GSILVGRALAPIEGAVGQWKNYIGARGAMDRLQTMLEEE--KSADHMPLPPEPRGVLSA 332
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 283 fvygylellfgrlrrlvasftltqsfasmdrvfqlidedyglkngvgagpieikgtrldi 342
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 333 EAASILPFGAQQPTMRQASFRIDAGAAVALVPSNAGKSSLLRGIVGVWPCACAGYIRLDG 392
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 343 dhvafgyndneapilklnistiekgetvafvmsggsktllnlprrfydvtsqqlldg 402
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 393 YDIKQWPEKLGRIHVGVLPODIELFSGTVAQNI--ARFEPESOEVIKATLAGVHEMIQ 450
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 403 hnldkftlgslnrqdglvgqdnllfstlvenlllgrptacd-eevveaakmanahdflm 461
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 451 SLPMGYDTAIGEGASISGGQRRLAARAVFRMPALLVLDENASLDQVGEVALMEAMK 510
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 462 nlpgydtvevgergvklsggqkgrlsariflfnppllldatsaldlesesllgeald 521
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 511 RLKAARTRVITATKVNLLAQAQDIYIMVINOQVISDFG 547
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 522 vl-skdrllivahrllstlhackivlenghivetyg 557
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: July 19, 2002, 10:43:17
 Job time: 137 sec

Db 23 TVGIFTAFINLMLVPSIYMLQYDRLVPSRNETTLLMLTLMLGFMGMSLLEYRSWV 82
QY 85 LVNGGLKFDGVARDPIFKSVLSDTLNR-KGIGQAFRDMQVREPMTG-GLIAFCDAWPT 142
Db 83 VIRIGSOLDMRLNTRYTAAYEANLNKSSDAGOMLSDLTNLQFLTGLSALFAFFDAPWF 142
QY 143 PVEVIYSWMLHPFEGILAIACIIIFGLAVMNDATKNPIQMATMASIAQNDAGSTLRN 202
Db 143 PLYLVIFLFPNMLGLFALVGAALLLALAVINEVSKRPLGEASKLSIMSGNLASTMLRN 202
QY 203 AEVWKAMGMNGGLQARRARDEQVAMQAAASDAGAVMSGIFYFRITVOTLLIGGAYL 262
Db 203 AEVIEALGMLPNLKRFRFGLHQRFLNSQRLASERASVYITITFVRMSLQSLVGLGGWL 262
QY 263 AIDGKISAGAMISILVGRALPIEGAVGOMKNYIGARGAMDLQTLREKSSADHMP 322
Db 263 AIDGHTIPGMNIGSILMGRTLPIEQVINWKSYSAAKLSYGRVLLLETHPQRTGMS 322
QY 323 LPREGVLSAFAASILRPPGAO-OPTMRQASFRIDAGAAVALVGSAGKSSLLRGIVGW 381
Db 323 LPREGVLSVEGYATPFGSKGDAVLHNVSFAIQPGVGLIGPSAGSKSTLARLVGIW 382
QY 382 PCAAGVIRLDGYDIKQMDPEKLGHNHGYLPQDILFSGTVAQNTARTPEPSQVITAEAT 441
Db 383 PVSGIYRLDNADLYQNNKDELGYIGLPPQDILFAGTIAENIARNDIDSEKVIKAAK 442
QY 442 LAGVHEMIOSLPMGYDPAIGEGASLSGGORLALRAVFRMPALLVLDEPNASLDQVG 501
Db 443 LAGVHELILRPNGYDVSIGNGAGLSGGQKORIGLARALYGDPAVLVDEPNSLDDAG 502
QY 502 EVALMEAMKRLKAARKRYVIFATHKVNLLAQADYIMVINGVISDFG 547
Db 503 EKALNOAIMFLKQRNKTVLTHRTNLLSMTSKLLLVNGVNAFQ 548

RESULT 2
US-08-620-605D-2
; Sequence 2, Application US/08620605D
; Patent No. 5846811
; GENERAL INFORMATION:
; APPLICANT: SHIBATANI, TAKEJI
; APPLICANT: AKATSUKA, HIROYUKI
; APPLICANT: KAWAI, ERI
; TITLE OF INVENTION: GENE PARTICIPATING IN THE MECHANISM OF
; TITLE OF INVENTION: SECRETION OF ESTERASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/620,605D
; FILING DATE: 22-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SVENSSON, LEONARD R
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 0020-3955
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)-205-8000
; TELEFAX: (703)-205-8050
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 588 amino acids
; TYPE: amino acid

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-620-605D-2

Query Match 41.8%; Score 1173.5; DB 2; Length 588;
Best Local Similarity 44.7%; Pred. No. 5.9e-104;
Matches 235; Conservative 111; Mismatches 177; Indels 3; Gaps 3;

QY 25 TAVSFIFINILALVSPYMLQYDRLVTSRNTLIVICVFLVFLVGLLEARTQV 84
Db 23 TVGIFTAFINLMLVPSIYMLQYDRLVPSRNETTLLMLTLMLGFMGMSLLEYRSWV 82
QY 85 LVNGGLKFDGVARDPIFKSVLSDTLNR-KGIGQAFRDMQVREPMTG-GLIAFCDAWPT 142
Db 83 VIRIGSOLDMRLNTRYTAAYEANLNKSSDAGOMLSDLTNLQFLTGLSALFAFFDAPWF 142
QY 143 PVEVIYSWMLHPFEGILAIACIIIFGLAVMNDATKNPIQMATMASIAQNDAGSTLRN 202
Db 143 PLYLVIFLFPNMLGLFALVGAALLLALAVINEVSKRPLGEASKLSIMSGNLASTMLRN 202
QY 203 AEVWKAMGMNGGLQARRARDEQVAMQAAASDAGAVMSGIFYFRITVOTLLIGGAYL 262
Db 203 AEVIEALGMLPNLKRFRFGLHQRFLNSQRLASERASVYITITFVRMSLQSLVGLGGWL 262
QY 263 AIDGKISAGAMISILVGRALPIEGAVGOMKNYIGARGAMDLQTLREKSSADHMP 322
Db 263 AIDGHTIPGMNIGSILMGRTLPIEQVINWKSYSAAKLSYGRVLLLETHPQRTGMS 322
QY 323 LPREGVLSAFAASILRPPGAO-OPTMRQASFRIDAGAAVALVGSAGKSSLLRGIVGW 381
Db 323 LPREGVLSVEGYATPFGSKGDAVLHNVSFAIQPGVGLIGPSAGSKSTLARLVGIW 382
QY 382 PCAAGVIRLDGYDIKQMDPEKLGHNHGYLPQDILFSGTVAQNTARTPEPSQVITAEAT 441
Db 383 PVSGIYRLDNADLYQNNKDELGYIGLPPQDILFAGTIAENIARNDIDSEKVIKAAK 442
QY 442 LAGVHEMIOSLPMGYDPAIGEGASLSGGORLALRAVFRMPALLVLDEPNASLDQVG 501
Db 443 LAGVHELILRPNGYDVSIGNGAGLSGGQKORIGLARALYGDPAVLVDEPNSLDDAG 502
QY 502 EVALMEAMKRLKAARKRYVIFATHKVNLLAQADYIMVINGVISDFG 547
Db 503 EKALNOAIMFLKQRNKTVLTHRTNLLSMTSKLLLVNGVNAFQ 548

RESULT 3
US-08-772-270A-4
; Sequence 4, Application US/08772270A
; Patent No. 6019984
; GENERAL INFORMATION:
; APPLICANT: Machines, Janet
; APPLICANT: Ricciardi, Paul
; APPLICANT: Mallard, Bonnie
; APPLICANT: Rosendal, Soren
; TITLE OF INVENTION: NOVEL BACTERIAL PREPARATIONS, METHOD FOR
; TITLE OF INVENTION: PRODUCING SAME, AND THEIR USE AS VACCINES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bereskin & Parr
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/772,270A


```

1      FLING DATE: December 23, 1996
2      CLASSIFICATION: 424
3      ATTORNEY/AGENT INFORMATION:
4      NAME: Gravelle, Michelle
5      REGISTRATION NUMBER: 40,261
6      REFERENCE/DOCKET NUMBER: 6580-81
7      TELECOMMUNICATION INFORMATION:
8      TELEPHONE: (416) 364-7311
9      TELEFAX: (416) 361-1398
10     INFORMATION FOR SEQ ID NO: 4:
11     SEQUENCE CHARACTERISTICS:
12     LENGTH: 707 amino acids
13     TYPE: amino acid
14     STRANDEDNESS: single
15     TOPOLOGY: linear
16     MOLECULE TYPE: protein
17     ORIGINAL SOURCE:
18     ORGANISM: Actinobacillus pleuropneumoniae
19     OS-08-772-270A-4

```

| | | | | |
|-----------------------|------------------|-------------------|-----------|------------|
| Query Match | 20.3% | Score 569.5 | DB 3 | Length 707 |
| Best Local Similarity | 28.4% | Pred. NO. 5.9e-46 | | |
| Matches 161 | Conservative 105 | Mismatches 253 | Indels 47 | Gaps 9 |

| | | | |
|----|-----|--|-----|
| QY | 10 | TIDQDAVBARPAVITAMVSPFINIALATSPYMLQVYRVSIVTSRNVSLTIVLTVCF | 69 |
| Db | 141 | TWFIIPAVIKTRKPIETFLIIVSIFLQIFALITPLEFQVVMKVLVHREFSLANTIVTAL | 200 |
| QY | 70 | LELVYGLLELRTOVLVRGBKEDGVARDDIFKSVSDSTLS-----RKIGGGAQFRMDQ | 124 |
| Db | 201 | VLFELIYINGLIRYIFEAHSTSRIDVELGARLFRHLLALPISEYENRRVSGTVARVELDQ | 260 |
| QY | 125 | VREPMTC-GLIARODAPWTEVEVIVSMHLRPFBSILAIACIIIFGLAVMNDNAKTNP | 163 |
| Db | 261 | IRNFLTQALTSVLDMFSTFPAVMVYSPKTLTVLIGSLPFYMGWSIFISPILRRLD | 320 |
| QY | 184 | MATMASTIAONDAGSTLRNKEVMMKAMOMGGLQARHARDEQVAVOAOAA-----SDAGGA | 239 |
| Db | 321 | EKPARGADNOSFLVESTATINTIKALAVTQMTW-----DKOLASVYSAGFRVTTIATI | 376 |
| QY | 240 | VMSGIKVERNIVOTLLIGGAYLAIDKISAGAMIASGILVGRALADIEGAVGOMKNYIG | 299 |
| Db | 377 | GOQGVJCIQKVMYVITLMLAHVLSGDSLIGQILARNMLSGQVIAYVIRLAQLOMQFOQ | 436 |
| QY | 300 | ARGAMDLOQTMKRE-EKSADDHMLRPPRQVLSAEASILRPGAQOPTMQASFR----- | 353 |
| Db | 437 | VGISVTRLGDVLSNPSTESYOKLALRPIKDDI-----TFRNIRFRYKPA | 481 |
| QY | 354 | -----IDAGAAVALVGPSPASSLSLRGIYVGPCAGVIRLBDGDIOMQDEKL | 403 |
| Db | 482 | PVILINDVNLSTOQGEVYIGVGRSGSGSTLTKLQRFYIPENQGVLLDGDHLLADLADNML | 541 |
| QY | 404 | GRHVGVLPDIDELFSGVAONIRAFTE-FESQEYIEATAGVHEMIOSSPMGDTAIGE | 462 |
| Db | 542 | RROGVVLDQNVLLNRSIRBNIALADGGMEMKIVHAAKLAGAHEFISELREGNTIYGE | 601 |
| QY | 463 | GGASLSGGORLALARAVFRMPALVLDPEPNASLDQVGEVALMEAMKRLKAARKVIFA | 522 |
| Db | 602 | QGAGLSGGORLAIARALVNNPKILFIEDATSLADYESEHIIMRNHQI-CKGRFTVILI | 660 |
| QY | 523 | THKVNNLAADYIMVINOVIISDPGE | 548 |
| Db | 661 | AHRLSITVKNADRIIVMKGQIVDQK | 686 |

RESULT 4
US-08-772-270A-12
; Sequence 12, Application US/08772270A
; Patent No. 6015984
; GENERAL INFORMATION:
; APPLICANT: Macinnes, Janet
; APPLICANT: Riccietatti, Paul

```

1  APPLICANT: Mallard, Bonnie
2  APPLICANT: Rosendal, Sven
3  TITLE OF INVENTION: NOVEL BACTERIAL PREPARATIONS, METHOD FOR
4  TITLE OF INVENTION: PRODUCING SAME, AND THEIR USE AS VACCINES
5  NUMBER OF SEQUENCES: 14
6  CORRESPONDENCE ADDRESS:
7  ADDRESSEE: Bereskin & Parr
8  STREET: 40 King Street West
9  CITY: Toronto
10 STATE: Ontario
11 COUNTRY: Canada
12 ZIP: M5H 3Y2
13
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Floppy disk
16 COMPUTER: IBM PC compatible
17 OPERATING SYSTEM: PC-DOS/MS-DOS
18 SOFTWARE: PatentIn Release #1.0, Version #1.30
19
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER: US-08/772,270A
22 FILING DATE: December 23, 1996
23 CLASSIFICATION: 424
24
25 ATTORNEY/AGENT INFORMATION:
26 NAME: Gravelle, Michelle
27 REGISTRATION NUMBER: 40,261
28 REFERENCE/DOCKET NUMBER: 6580-81
29 TELECOMMUNICATION INFORMATION:
30 TELEPHONE: (416) 364-7311
31 TELEFAX: (416) 361-1398
32
33 INFORMATION FOR SEQ ID NO: 12:
34 SEQUENCE CHARACTERISTICS:
35 LENGTH: 711 amino acids
36 TYPE: amino acid
37 STRANDEDNESS: single
38 TOPOLOGY: linear
39 MOLECULE TYPE: protein
40 ORIGINAL SOURCE:
41 ORGANISM: Actinobacillus pleuropneumoniae
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
9
```

| | | | | |
|-----------------------|-------------------|--------------------|------------|-------------|
| Query Match | 19.7%; | Score 552.5; | DB 3; | Length 711; |
| Best Local Similarity | 28.9%; | Pred. No. 2.5e-44; | | |
| Matches 165; | Conservative 117; | Mismatches 238; | Indels 51; | Gaps 14 |

[illegible]

Db 544 WLRQGVVLDQDVLNRSIRDNIALTDPBSMSEKVIYAAKLAGADHDFISELEGNTIY 603
QY 461 GEGGASISGGORORLALARAFVFMALLVDEPNASLDQGEVALMEAMRLKAARTVI 520
Db 604 GELGAGLSGGORRIARLALVNNPRILIFDEATFALDSESHIINQNMOKI-CHGTVI 662
QY 521 FATHVNLAAQADYIMVINGVISDFEGRD 551
Db 663 IIAHRLSTVKNADRIITIMERGHIVEOGKHNO 693

RESULT 5
US-08-895-522-3
; Sequence 3, Application US/08895522
; Patent No. 5858719
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Shah, Puryl
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN ATP-BINDING CASSETTE
; TITLE OF INVENTION: TRANSPORT PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/895.522
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0336 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 694 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Genbank
; CLONE: 1167982
; US-08-895-522-3

Query Match 14.3%; Score 401.5; DB 2; Length 694;
Best Local Similarity 25.1%; Pred. No. 7.2e-30;
Matches 143; Conservative 113; Mismatches 233; Indels 81; Gaps 19;

QY 23 VITANVSEFFITIALVSEPLVLDQYDVLSRNVSTLIVTVCFLEFLVGLLEALRT 82
Db 97 IIVPEMFYVADVSLNQMSG-NMLNLSDAIPVATMATAVLIG-----YG----- 139
QY 83 QVIVAGGLKFGVAVDPVLFKSVLDSTLSKKGIGGQAF-----RMDQVREFWVGGLIAC 137
Db 140 --VSRAGAAFFNEVKNVAVGKVAQNSIRR--IAKNVFLHLHLNDLGFHLSTGTGALSRAI 195

QY 138 DAPWTEPVIVS---WMLHPFEIILAIACILIEGLAVNNDNAFKPMIOMATWASIA--- 191
Db 196 DRGTGRISEFLASLVNLLPIYFEMMLVSSVLTKCGAOPALVLTGLGATFTFYAVIR 255
QY 192 -----AONDAGS-----TLRNAEYKAMGMWGGLQARRRAR-DEQVAMQAAAD 235
Db 256 WRTFRLENNKADNDAGNAIDSLNVEYKYFN-----NEXEADRYDGFLETFEASL 310
QY 236 AGGAVMSGIKVFNIVQTLILGGAVLADGKISAGAMTAGSLVGRALAPITGAVQMK 295
Db 311 KSTSLAMLNFGONALFSVGLTIAIVLASQ-IVAGLTVGDLYM-----VNGLLFQLS 363
QY 296 ---NYIG-----ARGA---WDRLOTMLREKSDDHMPLEPGGVLSAEASILPGAQ- 343
Db 364 LPLNFGTVYRETRQALIDMNTLFTLKVTRIKDKYMAPPLQ--ITPQATYAFDNVHF 421
QY 344 -----OPTMQASFRIIDAGAAVALVGPSAGKSSLLNGIYVWPCAGVIRLDGYDKQW 398
Db 422 EYIEGQKVLNGVSEFVYAGKKAVALVGGSGSKSTIVRLFRFEPOKGSYLAGQNLQDV 481
QY 399 DPEKGRHVGLPDQLELFSGYVAQNIARFT-EFESQVTEATLGVHEMISLPMGYD 457
Db 482 SLESLRAVGVVPODAVLFNHTIYVNLVGNINASPEEYVAVAKLAGLHDAIIRMPGYD 541
QY 458 TAIGEGASISGGORORLALARAFVFMALLVDEPNASLDQGEVALMEAMRLKAAR 517
Db 542 TVGERGLKLSGGEKQVAVARALTKNPVYLDEATSSLDSTIEETILGAMRV-VKHR 600
QY 518 TVIFATHKVNLAAQADYIMVINGVISDFG 547
Db 601 TSIFIAHRLSTVVDADIELVLSGKVAERG 630

RESULT 6
US-09-195-391-3
; Sequence 3, Application US/09195391
; Patent No. 6080842
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Shah, Puryl
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN ATP-BINDING CASSETTE
; TITLE OF INVENTION: TRANSPORT PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/195.391
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/895.522
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0336 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:

```
SEQUENCE CHARACTERISTICS:
LENGTH: 694 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1167982
US-09-195-391-3

Query Match      14.38; Score 401.5; DB 3; Length 694;
Best Local Similarity 25.18; Pred. No. 7.2e-30;
Matches 143; Conservative 113; Mismatches 233; Indels 81; Gaps 19;

QY 23 VITAMVSEFFNITLALSPYLMLOYDVLFSRWVSTLIVTCVFLVYGLLEALRT 82
   :::::::::::::::::::::
DB 97 IVFPMKRYAVDSLQMSG--NMLNLSDPNTVATMATRAVLIG-----YG----- 139
   :::::::::::::::::::::
QY 83 QVLVRGGLKFDGVARDPYFKSVLDSTLSRKIGGOAF-----RDMDOVERFMTGLIAFC 137
   :::::::::::::::::::::
DB 140 --VSRAGAAFENEVRNNAFEGKVAQNSIR--IAKNVFLHLNLDLGFHLSRQTALSKAI 195
   :::::::::::::::::::::
QY 138 DAPPTPVFVIVS--WMLHPEFGLIATIIIFGLAVMNDATKPIQATMASIA-- 191
   :::::::::::::::::::::
DB 196 DRGTGISFVLSALVFNLLPIVFEMMLVSVLYYKCGAFALVTGLGTAVTAFTAVTR 255
   :::::::::::::::::::::
QY 192 -----AONDAGS---TLRNAEVMKAMGMGGLQARMRAR--DEGVANQAAASD 235
   :::::::::::::::::::::
DB 256 WRTFRRIEMNKADNAGNAIDSLNLETYVYFN-----NEKYEAORDGLKTYETASL 310
   :::::::::::::::::::::
QY 236 AGGAVMSGIKVFNIVOTLLIGGAYLAIDKISAGAMIAGSIIVGRALPIEGAVGOMK 295
   :::::::::::::::::::::
DB 311 KSTSTLMLNCGAIAISVGLTAIMVLASQS--IVAGALTYDDVM-----VNLLEQLS 363
   :::::::::::::::::::::
QY 296 ---NYIG----ARGA---WDRLQTMREEKSADHMPLEPRGVLSAASILPPGAQ-- 343
   :::::::::::::::::::::
DB 364 LPLNLTGVYRETQALIDMTLTLLKVDTRIKDKVAPPLQ--ITPQTAVAFDNVHF 421
   :::::::::::::::::::::
QY 344 -----OPTMQASFRIAGAAVALVGPSAGKSSLLRGIVVWPCAGVIRLDGDIKOW 398
   :::::::::::::::::::::
DB 422 EYIGQVYLVNVSFEVPRKVAIVGSGSKSTIVLFFETPPQKSIYLAQONQOV 481
   :::::::::::::::::::::
QY 399 DPEKLGHHVGLIPODIELFSGTVAQNIARFT--EFSEQVIEAATLAGVHEMIQSLPMGYD 457
   :::::::::::::::::::::
DB 482 SLESIRAVGVYPODAVLFHNTIYYNLLYGNINSPEEVYAVAKLAGHDAILLMPHCYD 541
   :::::::::::::::::::::
QY 458 TAIGEGASISGGOROLAIARAVERMPALLIVDEPNASLDQVEVALMEAMKRLKAAR 517
   :::::::::::::::::::::
DB 542 TOYGERGLKLSGGEKQVAIAIRAILKNPVLVDEATSSIDSITEETILGAMRDV--VKHR 600
   :::::::::::::::::::::
QY 518 TVIFATHKVNLNQAADYIMVINOVISDFG 547
   :::::::::::::::::::::
DB 601 TSIFIAHRLSTVVDADDELIVLSGKVAERG 630
   :::::::::::::::::::::

RESULT
7
US-08-895-522-1
; Sequence 1, Application US/08895522
; Patent No. 5858719
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Shah, Puri
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN ATP-BINDING CASSETTE
; TITLE OF INVENTION: TRANSPORT PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
```

```
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/895,522
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PP-0336 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 747 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: OVARNOT02
CLONE: 545981
US-08-895-522-1

Query Match      14.38; Score 401.5; DB 2; Length 747;
Best Local Similarity 25.48; Pred. No. 8.1e-30;
Matches 150; Conservative 99; Mismatches 220; Indels 121; Gaps 21;

QY 23 VITAMVSEFFNITLALSPYLMLOYDVLFSRWVSTLIVTCVFLVYGLLEALRT 82
   :::::::::::::::::::::
DB 150 IVFPMKRYAVDSLQMSG--NMLNLSDPNTVATMATRAVLIG-----YG----- 192
   :::::::::::::::::::::
QY 83 QVLVRGGLKFDGVARDPYFKSVLDSTLSRKIGGOAF-----RDMDOVERFMTGLIAFC 137
   :::::::::::::::::::::
DB 193 --VSRAGAAFENEVRNNAFEGKVAQNSIR--IAKNVFLHLNLDLGFHLSRQTALSKAI 248
   :::::::::::::::::::::
QY 138 DAPPTPVFVIVS--WMLHPEFGLIATIIIFGLAVMNDATKPIQATMASIAQONDA 196
   :::::::::::::::::::::
DB 249 DRGTGISFVLSALV--FNLLPTIMEFVMLVSVLYYKCGAFALVTGLGTAVTAFTVA 305
   :::::::::::::::::::::
QY 197 GSTLRNAEVMKAMGMGGLQARMRARDEGVANQAAASDAGAVMSG-----IKVFRN- 249
   :::::::::::::::::::::
DB 306 -----VTRMRT--FRIEMNKADNAGNAIDSLNLETYKYRNE 344
   :::::::::::::::::::::
QY 250 -----IVOTLLIGGAYLAI--DGKISAG-----AMIASILVGR 282
   :::::::::::::::::::::
DB 345 RYEAQRYDGLKTYETASLSKSTLMLNFGQSAIFSGLTAIMVLASQGIIVACTGLTVG- 403
   :::::::::::::::::::::
QY 283 ALAPIEGAVGOMK---NYIG----ARGA---WDRLQTMREEKSADHMPLEPRGVLS 331
   :::::::::::::::::::::
DB 404 DLVAVNGILLQLSLPLNLTGVYRETQALIDMTLTLLKVDTRIKDKVAPPLQ--ITPQTAVAFDNVHF 454
   :::::::::::::::::::::
QY 332 AEAASILPPGAQ-----OPTMQASFRIAGAAVALVGPSAGKSSLLRGIV 378
   :::::::::::::::::::::
DB 455 ASPIQITPQATVAFVFNWHEFYIGQAVLSISFEVARGKVAIVGSGSKSTIVALLF 514
   :::::::::::::::::::::
QY 379 GWPCAGVIRLDGYDIKQMDPEKLGHHVGLIPODIELFSGTVAQNIARFTFEFS--QEVY 437
   :::::::::::::::::::::
DB 515 RFEYPPQKSIYLAQONIDVLSLESIRAVGVYPODAVLFHNTIYYNLLYGNINSAPREV 574
   :::::::::::::::::::::
QY 438 EAATLAGVHEMIQSLPMGYDPAIGEGASISGGOROLAIARAVERMPALLIVDEPNASL 497
   :::::::::::::::::::::
DB 575 AVAKLAGHDAILLMPHCYDTPQYGERGLKLSGGEKQVAIAIRAILKNPVLVDEATSS 634
   :::::::::::::::::::::
```


QY 511 -RLKAKRVIETATHKVNLIAQADYIMVINGVISDFGE 548
Db 621 DNFTSGRTSVYIAHRLTIADADKIIVLNGRVEGK 659

RESULT 11

US-08-784-649A-2
; Sequence 2, Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
; APPLICANT: Sikic, Branimir I
; APPLICANT: Chen, Gang
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
; TITLE OF INVENTION: CYCLOSPORIN MODULATION
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/784,649A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: Reg No. 5830697 36,677
; REFERENCE/DOCKET NUMBER: 06037/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1279 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-784-649A-2

Query Match 13.5%; Score 379.5; DB 2; Length 1279;
Best local similarity 24.6%; Pred. No. 2,4e-27;
Matches 142; Conservative 99; Mismatches 230; Indels 107; Gaps 18;

QY 26 AMVTSFFINILALVSPLYMLQVDRVLSRVSTL-----IVLVICVFLVLYVGLLEAL 80
Db 728 AITPSKIIGVETRI-----DDPEFKRONSLFSLFLALGIIISFTFFLOGFTFGK 778
QY 81 RTQVLVKGGLKFDGVDARDPIKSVL-----DSTLSRKGIQGO 117
Db 779 AGEILITKR-----LRYMFRSMLRQDVSWFDDPKNTGALTTRLANDAQVGAIGSR 831
QY 118 AFRMDQVREFMTGLIAFCDAPTVPVFIYSWMLHPFFGILAIACIIIFG---LAVYN 174
Db 832 LAVITONIANIGTGIISF-----IYGMQUTLL--LLAIVPIAIVAGVEMKMLS 879
QY 175 DNATKNPIQAMATMASIAQ--NDAGSTLRNAEYKAKGMG--GIQARRARDEQVAMQA 231
Db 880 GQALKKKELEGAGKIKATEALENRTVSLTQEQKFEHMYAQSLQVPR-----928
QY 232 AADGAGVAMSGIKV--FRNIVQTLILG---GAYLAI-----DGKISAGAMIAGSIL 279
Db 929 --NSLRRAHITGITFTQAMMTFSYAGCFRFGAIVLAHKLMSEFDVLLVSAVYFGAMA 986

QY 280 VGR--ALAPIEGAVGQKNITGARGANDRL-----QTMLEEKSSADDHMLPPEPRGLSA 332
Db 987 VGOVSSFP-----DAKAKISAHIIMIEKTPLDISTYSTEGLMNTLEGNTFG 1037
QY 333 EAASILPQGAQOPTMROASRIDAGAVALVGPSAAGSSILRGIVGVMPCCAGVIRLDG 392
Db 1038 EYVFNYPTRPDIPLVQLSLVKKGGTLLVGSAGCKSTVYVOLLSEFFYPLACKVILDG 1097
QY 393 YDIKQWPEKLRHVGYLPDIELEFSGTVAONIA--RTEFESQEVIEATLAGVHEMT 449
Db 1098 KEIKRLNWQMLRAHLGIQSEPIIFDCSIAENIAYGDSRVSOEIVRAAKEANIHAFI 1157
QY 450 QSLPMGYDTAIGEGASLSGQQRALARAVPFMPALLVDEPNASLDQVGEYALMEAM 509
Db 1158 ESLPNKYSTVWGKGTQSGQKORAIARALVQPHILLDEATSLDRESEKVOEAL 1217
QY 510 RLKAKRVIETATHKVNLIAQADYIMVINGVISDFG 547
Db 1218 DKAREG-RTCIIVAHRLSTIONADLIYVFONGRVEKENG 1254

RESULT 12

US-08-583-276-19
; Sequence 19, Application US/08583276
; Patent No. 5837536
; GENERAL INFORMATION:
; APPLICANT: McDonagh, Kevin T.
; APPLICANT: Nienhuis, Arthur
; APPLICANT: Tolstoshev, Paul
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN
; TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED
; SELECTION OF CELLS TRANSDUCED WITH SUCH GENES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cecchi & Stewart
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: DM4.V2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,276
; FILING DATE: 05-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/332,444
; FILING DATE: 31-OCT-1994
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1280 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-583-276-19

Query Match 13.5%; Score 379.5; DB 2; Length 1280;
Best local similarity 24.6%; Pred. No. 2,4e-27;
Matches 142; Conservative 99; Mismatches 230; Indels 107; Gaps 18;

QY 26 AMVTSFFINILALVSPLYMLQVDRVLSRVSTL-----IVLVICVFLVLYVGLLEAL 80
Db 729 AITPSKIIGVETRI-----DDPEFKRONSLFSLFLALGIIISFTFFLOGFTFGK 779

```

OY 81 RTQVLVRLGKFDGVARDPFEKSVL-----DSTLSRKIGGQ 117
   :|:|:|
Db 780 AGEILTKR-----LRIMVFRSMLRQDVSWFDDPKNTGALTTLRLANDAOVKATISR 832
OY 118 AFRDMDOVEEMTGLIAFCDAWPTPVFIVSWMLHPFEGILATIIACIIIFG---LAVMN 174
   :|:|:|
Db 833 LAVIQNIANLGTGIIISF-----IYGQWLTLL--LLAIPIIIAIGVEMKMLS 880
OY 175 DNATNPITQMATMASIAA--NDAGSTLRNAEVMKAMGMG--GLQARRARARDEYANQA 231
   :|:|:|
Db 881 GOALKDKKLELGAGKIAIEAIEENFTVVSLTQEOKEFHMVQASLOVPR-----929
OY 232 AASDAGVMSGIRK-FNNIVQTLILG---GAYLAI-----DGKISAGMIAGSIL 279
   :|:|:|
Db 930 --NSLRKAHIGITFSFQAMMYFSYACFRGATLVAAHKLMSEFDVLIVSAVFGAMA 987
OY 280 VGR--ALAPIEGAVGMKNYIGARGAMDRL-----QTMLEEKSSADHMPLEPRGVLSA 332
   :|:|:|
Db 988 VGOVSSFP-----DYAKAKISAHHIIMITEKTPIDISYTEGLMPTLEGVTFEG 1038
OY 333 EAASILPPGAQOPTMROASFRIDAGANALVGPSAGKSSLLRGIVGWPACAGYIRLDG 392
   :|:|:|
Db 1039 EVFENYPRPDIPLYQGLSLEVKKGQTLALVSSCGKSTVQVQLERFYDPLAGKVLIDG 1098
OY 393 YDIKOMDEKLGRIHGYLPDQIELEFSGTVAONIA---RTEFESEQVTEAATLAGVHEMI 449
   :|:|:|
Db 1099 KEIKRLNQWMLRAHGIYSQEPILFDCSIAERIAVAGNSRVVSOEIVRAKAEINIIHAFI 1158
OY 450 QSLPMGYDTAIGEGASLSGQORLALRAVFRMPALLVDEPNASLDQGEVALMEAM 509
   :|:|:|
Db 1159 ESLPKYSTKVGDKGTQLSGGOKRIALRALVRPHILLDEATSALDTESEKVVQEAL 1218
OY 510 KRLKAARTVIFATIKVLLAQAQDIYIMYINQVSDFG 547
   :|:|:|
Db 1219 DKAREG-RTCIIVIAHRLSTIQNADLIVFONGRVKEHG 1255

RESULT 13
5206352-4
: Patent No. 5206352
: APPLICANT: Roninson, Igor B.; Pastan Ira H.; Gottesman,
: Michael M.
: TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA
: SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
: NUMBER OF SEQUENCES: 4
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/622,836
: FILING DATE: 24-SEP-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 892,575
: FILING DATE: 01-AUG-1986
: APPLICATION NUMBER: 845,610
: FILING DATE: 28-MAR-1986
: SEQ ID NO: 4
: LENGTH: 1280
5206352-4

Query Match 13.5%; Score 379.5; DB 6; Length 1280;
Best Local Similarity 24.6%; Pred. No. 2.4e-27;
Matches 142; Conservative 99; Mismatches 230; Indels 107; Gaps 18;

```

```

Db 833 LAVIQNIANLGTGIIISF-----IYGQWLTLL--LLAIPIIIAIGVEMKMLS 880
OY 175 DNATNPITQMATMASIAA--NDAGSTLRNAEVMKAMGMG--GLQARRARARDEYANQA 231
   :|:|:|
Db 881 GOALKDKKLELGAGKIAIEAIEENFTVVSLTQEOKEFHMVQASLOVPR-----929
OY 232 AASDAGVMSGIRK-FNNIVQTLILG---GAYLAI-----DGKISAGMIAGSIL 279
   :|:|:|
Db 930 --NSLRKAHIGITFSFQAMMYFSYACFRGATLVAAHKLMSEFDVLIVSAVFGAMA 987
OY 280 VGR--ALAPIEGAVGMKNYIGARGAMDRL-----QTMLEEKSSADHMPLEPRGVLSA 332
   :|:|:|
Db 988 VGOVSSFP-----DYAKAKISAHHIIMITEKTPIDISYTEGLMPTLEGVTFEG 1038
OY 333 EAASILPPGAQOPTMROASFRIDAGANALVGPSAGKSSLLRGIVGWPACAGYIRLDG 392
   :|:|:|
Db 1039 EVFENYPRPDIPLYQGLSLEVKKGQTLALVSSCGKSTVQVQLERFYDPLAGKVLIDG 1098
OY 393 YDIKOMDEKLGRIHGYLPDQIELEFSGTVAONIA---RTEFESEQVTEAATLAGVHEMI 449
   :|:|:|
Db 1099 KEIKRLNQWMLRAHGIYSQEPILFDCSIAERIAVAGNSRVVSOEIVRAKAEINIIHAFI 1158
OY 450 QSLPMGYDTAIGEGASLSGQORLALRAVFRMPALLVDEPNASLDQGEVALMEAM 509
   :|:|:|
Db 1159 ESLPKYSTKVGDKGTQLSGGOKRIALRALVRPHILLDEATSALDTESEKVVQEAL 1218
OY 510 KRLKAARTVIFATIKVLLAQAQDIYIMYINQVSDFG 547
   :|:|:|
Db 1219 DKAREG-RTCIIVIAHRLSTIQNADLIVFONGRVKEHG 1255

RESULT 14
US-08-752-447-2
: Sequence 2, Application US/08752447
: Patent No. 5994088
: GENERAL INFORMATION:
: APPLICANT: Mechtner, Eugene
: APPLICANT: Roninson, Igor B
: TITLE OF INVENTION: Methods and Reagents for Preparing and
: TITLE OF INVENTION: Using Immunological Agents Specific for P-glycoprotein
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: McDonnell Boehrnen Hulbert & Berghoff Ltd.
: STREET: 300 South Wacker Drive, Seventh Floor
: CITY: Chicago
: STATE: Illinois
: COUNTRY: USA
: ZIP: 60606
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/752,447
: FILING DATE: 15-NOV-1996
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: No. 5994088nan, Kevin E
: REGISTRATION NUMBER: 35,303
: REFERENCE/DOCKET NUMBER: 95,1121
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312-913-0001
: TELEFAX: 312-913-9808
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1280 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-752-447-2

```

Query Match 13.3%, Score 373.5; DB 2; Length 1280;
Best Local Similarity 24.6%, Pred. No. 8.9e-27;
Matches 142; Conservative 96; Mismatches 233; Indels 107; Gaps 18;

QY 26 AMVPSFFINILALVSP-----LYMLQYDRLVLSRNVSTLLVLYVCVLF-----LV 73
DB 729 AIFSKIIIGVFTRI-----DDPETRKONSLSLFLALGIIISFTFFLOGFTRGK 779
QY 81 RTQVLVRGGLKFDGVARDPFKSVL-----DSTLSRKIGGQ 117
DB 780 AGEILITR-----LRVNFERSMLRODVSFHPDKNTGALTTLRLANDAOVKALISGR 832
QY 118 AFRDMQVREEMTGLLAFCDAPWTPVEVYSWMLHPFGLIATIIIGLAVMN--- 174
DB 833 LAVIQTNIANTIGTIIISF-----IYQWOLTL--LLAIYPIIATAGVEMKKFA 880
QY 175 DNATNPQIOMATMASIAAO--NDAGSTLRNMEVKKAMGWS--GLQARRARRDEQVAMQA 231
DB 881 GOALDKKKELEGAGKIATEAIENFRTVVSLLTOEQKEHHWYASLOVPYK----- 929
QY 232 AASDAGAVMSGIKV-FRNIVQTLILG---GAYLAI-----DGKISAGAMIASIL 279
DB 930 --NSLRKMHIFGTFSPQAMWFSYAGCFRGATLVAKHKLMSFDVLLVFSAVFGAMA 987
QY 280 VGR--ALAPIEGAVGOMKNYIGARGAMDRL-----QTMIREKSSADHMPLEPRGVLSA 332
DB 988 VGVVSSFP-----DYAKAKISAHHIIMIEKTPIDBSYSTEGLMPTLEGNVTFG 1038
QY 333 EASLILPGGAQOPTMRQASFRIDAGAAVALVGPSAAGSSILRGIVGWPCAGVIRIDG 392
DB 1039 EGVENVYPRPDIPVLQGLSLEVKKQOTLALVSSGCGKSTVVQLERFIDPLAGVYLLDG 1098
QY 333 YDIKQDEPKELGRHVGYPQDIELEFSGTVACNIA--REFEFSEVIEAATLACVHEMI 449
DB 1099 KEIKRLANQWMLRAHGIYSQEPILFDCAIENIAGVDSNRVVSQGEIYRAAEANIHAFI 1158
QY 450 OSLEPGYDTAIGEGASISGGORQLALARAVERPALIVDEPNASLDQVEVALMEAM 509
DB 1159 ESLPKYSKRVKDGKTOISGGOKORIALRRLVROPHILLDEATSSALDTESEKVOEAL 1218
QY 510 KRLKAARTVIEFATKRVNLLAQADYIMVNOGVISDFG 547
DB 1219 DKAREG-RICIVIAHRLSTIONADLIVFQNGRVKHEG 1255

RESULT 15
US-08-232-537-2
Sequence 2, Application US/08232537
Patent No. 5516655
GENERAL INFORMATION:
APPLICANT: Peery, Robert B.
APPLICANT: Skatrud, Paul L.
TITLE OF INVENTION: MULTIPLE DRUG RESISTANCE GENE OF
TITLE OF INVENTION: AUREOBASIDIUM PULLULANS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,537
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas G.

REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X9212
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
TELEFAX: 317-276-1917
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1302 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-232-537-2

Query Match 13.3%, Score 373.5; DB 1; Length 1302;
Best Local Similarity 25.1%, Pred. No. 9.1e-27;
Matches 151; Conservative 106; Mismatches 205; Indels 139; Gaps 26;

QY 26 AMVPSFFINIL-ALVSP-----LYMLQYDRLVLSRNVSTLLVLYVCVLF-----LV 73
DB 740 AIVFENLISVLNDESPDERSRADLFSILF-----ILALIALFSYAGNGCC 786
QY 74 YGL-----LEALRTQVLVRGGLK-EDGVARDPFKSVLSDSTLSRKIGGQAFRDMQ 124
DB 787 FGIVSHFVAKIQHTSLASILRQDMQFSGQSVPSLMSLSDAGQLACLSVAIGTIFT 846
QY 125 VREFMTGCLIAFCDAWTPVEVYSWMLHPFGLIATIIIGLAVMNDATKNPIOM 184
DB 847 VCVSTIGTII-----LAHVVAWK-----IAVLLAAVPMV---ITAGVYRL 884
QY 185 ATMASI-----AAONDAQS-----TRNMEVKKAMGGLQARRARRDEQVAMQAASD 235
DB 885 RVLALAESRHRSAVNDASIAEACRGIRTTIASIGREGV-----SRASN 929
QY 226 AGGAVM--SGIKVERNIYOTLI-----LGGGAYLAIDGKISA 270
DB 930 A-AVKEPPDKIR-FTLITNTLLALSITYFYALAYWMCAGAKYRNQTSQDFFIVL 986
QY 271 GAMINGSLVGR--ALAPIEGAVGOMKNYIGAR--GAMDRLOTMLREKSSADHMP--- 322
DB 987 PALPSAGSAGQIFSLSPMSRAG-----VAARNVFGIHDQKPTIYVDARQSGALPES 1041
QY 323 -----LPEPRGVLSAEAAST-LPPGGAQOPTMRQASFRIDAGAAVALVGPSAAGKS 371
DB 1042 LSIPTLEDKASPSGSGWIEFKNVSCLYPSKQHPALQNVNISIRGEFTALVPSGAGKS 1101
QY 372 SLIRGIVGWPCAGVIRIDGYDIKQMD-PEKIGRHHVGLPQDIELEFSGTVACNIA--AR 427
DB 1102 TILSLQRYDPTAGSVQDGDIREVAVPOHRRG-LGIYPOEPDLPFGSISYNTIGLGA 1160
QY 428 FTEPESQEVIE-AATLAGVHEMIQSLPMGYDTAIGEGASISGGORQLALARAVERMPA 486
DB 1161 POLVLRDIDIEKICAKGCIHERFIMSLPEGSTECGNSKLSGGOKORIAVAFALIRPSE 1220
QY 487 LLVDEPNASLDQVEVALMEAMKRLKAARTVIEFATKRVNLLAQADYIMVNOGVISDF 546
DB 1221 VLLDEYISALDAHSEQQTKEAVDG-ASVDRTTIVVAHRLSTVQANADRLVFDGRVVEV 1279
QY 547 G 547
DB 1280 G 1280

Search completed: July 19, 2002, 10:41:36
Job time: 36 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2002, 10:41:00 ; Search time 32.83 seconds
(without alignments)
1624.416 Million cell updates/sec

Title: US-09-913-414-4
Perfect score: 2807
Sequence: 1 MFRSGAKPTIFDQAVLVAR.....MNINGVISDFGERDRCWPS 555

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|----------|---------------------|
| 1 | 2765 | 98.5 | 578 | 2 D87374 | RsaA secretion sys |
| 2 | 1231 | 43.9 | 593 | 2 S26696 | alkaline proteins |
| 3 | 1136 | 41.2 | 600 | 2 C83221 | transport proteins |
| 4 | 1077 | 38.4 | 582 | 2 D96003 | ABC transporter AT |
| 5 | 1065.5 | 38.0 | 602 | 2 AF0477 | ABC transporter pr |
| 6 | 1044.5 | 37.2 | 564 | 2 E70394 | ABC transporter (h |
| 7 | 1039 | 37.0 | 576 | 2 A49933 | proteinase SM tran |
| 8 | 1039 | 37.0 | 585 | 2 T48672 | ABC-type transport |
| 9 | 1029 | 36.7 | 615 | 2 E96310 | rhilobidin secret |
| 10 | 1029 | 36.7 | 615 | 2 AF2972 | alkaline proteinase |
| 11 | 1026.5 | 36.6 | 575 | 2 S12525 | metalloproteinase |
| 12 | 1012 | 36.1 | 583 | 2 E71687 | hypothetical prote |
| 13 | 1011 | 36.0 | 584 | 2 E97753 | hypothetical prote |
| 14 | 958 | 34.1 | 589 | 2 D95953 | hypothetical prote |
| 15 | 935 | 33.3 | 577 | 2 AH2908 | hypothetical prote |
| 16 | 935 | 33.3 | 588 | 2 G97683 | prsd protein (U891 |
| 17 | 931 | 33.2 | 690 | 2 D96318 | ABC protein AGR_L |
| 18 | 930 | 33.1 | 583 | 2 AI3098 | hypothetical prote |
| 19 | 930 | 33.1 | 606 | 2 H98187 | prsd protein (U891 |
| 20 | 927 | 33.0 | 551 | 2 AI2964 | hypothetical prote |
| 21 | 609 | 21.7 | 989 | 2 AE2140 | hypothetical prote |
| 22 | 607.5 | 21.6 | 1003 | 2 AE2335 | toxin secretion AB |
| 23 | 592.5 | 21.1 | 706 | 2 T00228 | hemolysin B - Esch |
| 24 | 584.5 | 20.8 | 707 | 2 AG1378 | leukotoxin express |
| 25 | 577.5 | 20.6 | 1011 | 2 S75806 | hemolysin secretio |
| 26 | 576.5 | 20.5 | 708 | 2 C30169 | leukotoxin express |
| 27 | 575.5 | 20.5 | 725 | 2 T03514 | probable secretion |
| 28 | 573.5 | 20.4 | 975 | 2 AC2517 | hypothetical prote |
| 29 | 572.5 | 20.4 | 742 | 2 H81855 | probable cytolysin |

| | | | | | |
|----|-------|------|-----|----------|---------------------|
| 30 | 569.5 | 20.3 | 707 | 2 D43599 | hemolysin I secret |
| 31 | 567.5 | 20.2 | 742 | 2 H81088 | ABC transporter fa |
| 32 | 562.5 | 20.0 | 707 | 1 LEEBBV | hemolysin secretio |
| 33 | 562.5 | 20.0 | 720 | 2 E82561 | toxin secretion AB |
| 34 | 561.5 | 20.0 | 707 | 2 S10057 | ABC-type transport |
| 35 | 559.5 | 19.9 | 707 | 1 LEECB | hemolysin secretio |
| 36 | 559 | 19.9 | 893 | 2 AH2007 | toxin secretion AB |
| 37 | 552.5 | 19.7 | 711 | 2 C49219 | toxin apxIII secre |
| 38 | 549.5 | 19.6 | 723 | 2 C83412 | probable ATP-bindl |
| 39 | 544 | 19.4 | 720 | 2 H82198 | RTX toxin transport |
| 40 | 543.5 | 19.4 | 704 | 2 H82381 | toxin secretion AT |
| 41 | 533 | 19.0 | 712 | 1 BVBRCB | cyab protein - Bor |
| 42 | 531.5 | 18.9 | 706 | 2 G90696 | hypothetical prote |
| 43 | 531.5 | 18.9 | 720 | 2 C85547 | probable cytoplasm |
| 44 | 522 | 18.6 | 712 | 2 AD0274 | probable toxin tra |
| 45 | 521.5 | 18.6 | 718 | 2 AF0835 | probable type I se |

ALIGNMENTS

| | | |
|---|--------|---|
| RESULT | 1 | |
| D87374 | | |
| RsaA secretion system, ATP-binding protein RsaD [imported] - Caulobacter crescentus | | |
| C:Species: Caulobacter crescentus | | |
| C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #extl_change 20-Apr-2001 | | |
| C:Accession: D87374 | | |
| R:Netman, W.C.; Feildblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gelin, M.L.; Haft, D.H.; Ko n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001 | | |
| A:Title: Complete Genome Sequence of Caulobacter crescentus. | | |
| A:Reference number: A87249; MGIID:21173698; PMID:11259647 | | |
| A:Accession: D87374 | | |
| A:Status: preliminary | | |
| A:Molecule type: DNA | | |
| A:Residues: 1-578 <STO> | | |
| A:Cross-references: GB:AE005673; NID:g13422298; PIDN:AAK22992.1; GSPDB:GM00148 | | |
| C:Genetics: | | |
| A:Gene: CC1008 | | |
| Query Match | 98.5%; | Score 2765; DB 2; Length 578; |
| Best Local Similarity | 99.8%; | Pred. No. 9.1e-189; |
| Matches 549; Conservative | 0; | Mismatches 1; Indels 0; Gaps 0; |
| QY | 1 | MFRSGAKPTIFDQAVLVARPAVITAMVSEFTINILALVSPLYMLQYDRLTSRNVSTL 60 |
| DB | 1 | MFRSGAKPTIFDQAVLVARPAVITAMVSEFTINILALVSPLYMLQYDRLTSRNVSTL 60 |
| QY | 61 | IVLTIVCVPLFLVYGLLEALRTQVLYRGGLKPDGVARDPIFKVLDSTLSRKIGGOARR 120 |
| DB | 61 | IVLTIVCVPLFLVYGLLEALRTQVLYRGGLKPDGVARDPIFKVLDSTLSRKIGGOARR 120 |
| QY | 121 | DMQOQVEFMFGILAFCDAPWTPVFIYSWMLHPFGIILAIICIIIFGLAVNNDATKN 180 |
| DB | 121 | DMQOQVEFMFGILAFCDAPWTPVFIYSWMLHPFGIILAIICIIIFGLAVNNDATKN 180 |
| QY | 181 | PIQMATWASIAONDAAGSTLRNAEVMKANGMWGGLQARWRARDEQVAVQAAASDAGAV 240 |
| DB | 181 | PIQMATWASIAONDAAGSTLRNAEVMKANGMWGGLQARWRARDEQVAVQAAASDAGAV 240 |
| QY | 241 | MSGIKVFNNIVQTLITGGGAYLAIDKISAGAMIASIILVGRALADIEAGVGMKNYICA 300 |
| DB | 241 | MSGIKVFNNIVQTLITGGGAYLAIDKISAGAMIASIILVGRALADIEAGVGMKNYICA 300 |
| QY | 301 | RCAMWDLQTMLEREKSAODHMPLEPRGVLSAFAASITLPPGAQOPTMROASPRIDGAAY 360 |
| DB | 301 | RCAMWDLQTMLEREKSAODHMPLEPRGVLSAFAASITLPPGAQOPTMROASPRIDGAAY 360 |
| QY | 361 | ALVGPAAKSSILRIGIVGVWPCACAVIRLDGYDKQMPDEKIGRRVGYLPDIDELFSGT 420 |
| DB | 361 | ALVGPAAKSSILRIGIVGVWPCACAVIRLDGYDKQMPDEKIGRRVGYLPDIDELFSGT 420 |

Db 440 AAARLGVHDLVRLPGGYDTRLDGAGLGGGCGORIGLALYGRPALIVDEPNASL 499
QY 498 DQGEVALMEAMKRLKAARTVIFATHKVNLQAQDIYIMINOGVISDFG 547
Db 500 DEAGEAALAEALAMRRGSSLVLVTHKPAVALTDLKLLHGGRIQRFQ 549

RESULT 4

ABC transporter ATP-binding protein prsd [imported] - Shnorhizobium meliloti (strain 102 D96003)
C:Species: Shnorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: D96003
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: D96003
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-582 <KUR>
A:Cross-references: GB:AL591985; PIDN:CAQ49692.1; PID:g15141179; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid pSymB
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, L.; Hyman, R.W.; Jones, T. Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, Hebut, F.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Welis, D.H.; Wong, K.; Yeh, K. A:Title: The composite genome of the legume symbiont Shnorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: prsd. Smb21466
A:Genome: plasmid

Query Match 38.4%; Score 1077; DB 2: Length 582;

Best local Similarity 42.2%; Pred. No. 7,9e-69; Matches 226; Conservative 104; Mismatches 202; Indels 4; Gaps 3;

QY 20 RPAVITAMVSEFINITIALVSPYLMQYDVLTSRNVSTLVIVYCVFLFVYGLLEA 79
Db 19 RAAFIGVAVSALVNLVLTGSEFMLEVDRLIPRSSTPILSLALLLVAFQAFEL 78
QY 80 LKQVVLVREGKLDGVARDPFEKSVLDSTLS--RKGIQGAQFDMQVREFMTG-GLIAR 136
Db 79 IRGRMLVRIAGALDESLNGRIYRAIVKAPLKLMOGDGLQALRDPQVRSFLSGVGPAAL 138
QY 137 CDAPWTFVETVYSWMLHPFEGILAITACITIFGLAVMNDATKNPLOMATMASIAQND 196
Db 139 FDLPRMPLFYAICFLFRHYVIGLAIIGGLITLITLTNGTQAPAKKSEAGCLRNVFA 198
QY 197 GSTLRNAEYKAKMGWGLQARRRARDEQVAAQAAADGAGVMSIGKFRNIYOTLIL 256
Db 199 OASORNAEYVHAAGMSARLTALWERTNTERFRDNRTSDINGYGALSKVFRALQSVL 258
QY 257 GGAAYTALDCKISAGAMIASIIVGRALAPIEGAVGOMKATYIGARGAMRLQTMLEKES 316
Db 259 AAGAVLVIRKESAPGIIIGSITLARALPVELAIGNMRLVNAAROSWOLKELNALPE 318
QY 317 ADQHMPLPEPRGVLSAEASILPRGAOOPTMQASFRIDGAVVALVGPSPAAKSSILRG 376
Db 319 ADAPRLQPRPHENLVEGLASGRPAORLVSVNFTVRAAGVAVYIGSASAKSSILARA 378
QY 377 YGVWPCAGAVIRLDGYDIKQMDPEKIGRHVGYLPDITELFSGTVAQNTARTE-ESQE 435
Db 379 ILGIWPAVYSGSVRLDGAALDQMSDALGKHVGYLPQDVLEFAGTANICRIFEDATSEA 438
QY 436 VIEAATLAGVHEHIOQLPMGYDTAIEGGASLSGGORQRLALARAVERMPALLVDEPNA 495
Db 439 IVAAARAAARVNDILRLPNGYDTEIGDGWTLISAGORORVALARALVGGPFLVLDPEPS 498

QY 496 SLDOGEVALMEAMKRLKAARTVIFATHKVNLQAQDIYIMINOGVISDFGERD 551
Db 499 NLDAEGQALSEALMSVRSRGIVYVAHRPSALASDVLMMEGMAQFPRKEQ 554

RESULT 5

ABC transporter protein YPO3920 [imported] - Yersinia pestis (strain C092) AF0477
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C:Accession: AF0477
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M delo-Terrage, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G L.M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AF0477
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-602 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAQ93386.1; PID:g15981832; GSPDB:GN00175
A:Genetics:
A:Gene: YPO3920

Query Match 38.0%; Score 1065.5; DB 2: Length 602;

Best local Similarity 41.5%; Pred. No. 5,4e-68; Matches 227; Conservative 108; Mismatches 203; Indels 9; Gaps 5;

QY 5 SGAKRTFDQAVVAVRPAVITAM-VFSFINITIALVSPYLMQYDVLTSRNVSTIVL 63
Db 10 SSVPTTL--SVLAGNKKIIMGICLFTAVINLMLAPAIMLOVYDVLASANTMTLML 67
QY 64 TVICVFLFVYGLLEALRTQVLYVREGKLDGVARDPFEKSVLDSTLSRK--IGQAFR 120
Db 68 TVLVGVFEVFIQLEWVRSVAVIRLTQIDMQNPENAFNAANL--KGNHNPAAQALN 125
QY 121 DMOQVREFMTG-GLIARCDAPWTFVETVYSWMLHPFEGILAITACITIFGLAVMNDATK 179
Db 126 DLTVLRQFATGNALFAFFDAPWPEFLVLLVFLPWLGMGLAAAGAILLVYLAWLNMWICK 185
QY 180 NPLOMATMASIAQNDAGSTLRNAEYKAKMGWGLQARRRARDEQVAAQAAADGAGA 239
Db 186 KPLHDASITTSHTQOANANLRNADYIEAMGMLKALBERMLQHANFLYQONLASDKSSR 245
QY 240 VMSGKFRNIYOTLILGGAYTALDCKISAGAMIASIIVGRALAPIEGAVGOMKATYIG 299
Db 246 VFAVAKSSRQALQSMMLGGLALLVYNEIRFAGVMIAGSILIGRVLPIDQLVIMWQMSH 305
QY 300 ARGAMDRLOTMLEKESADDMPLPEPRGVLSAEASILPRGAOOPTMQASFRIDAGAA 359
Db 306 ARLAYORLSQLAQHPSSPGWVLPAPQGLNVTOLMACRPGTHIPVLSINFELOPGDV 365
QY 360 VALVGPSPAAKSSILKRIYGVWPCAGAVIRLDGYDIKQMDPEKIGRHVGYLPDITELFSG 419
Db 366 LGTIGFSGSKSTLAKILVASQPTFSGTVRLDSADLSRMDKTQGLERIGLIPONTQIFRG 425
QY 420 TVANQNTARTEFESOEYEAATLAGVHEHIOQLPMGYDTAIEGGASLSGGORQRLALAR 479
Db 426 TVENATARFCAIDTAKVNAAGLADVHDLHLPGYDTPLDGDSGLSGGQRITALAR 485
QY 480 AVYRMPALLVDEPNASLDQGEVALMEAMKRLKAARTVIFATHKVNLQAQDIYIMIN 539
Db 486 AMYGIPLRILVDEPNASLDKEGQALLASTIQKQOGCTIVMTTHRPBELLSGSDYLLFLK 545
QY 540 QGVISDF 546
Db 546 NGQMDLP 552

RESULT 6

E70394
ABC transporter (hlyB subfamily) - Aquifex aeolicus
A:Title: Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 02-Feb-2001
C:Accession: E70394
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:9819666
A:Accession: E70394
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-564 <AOF>
A:Cross-references: GB:AE000723; NID:g2983569; PIDN:AAC07149.1; PID:g2983577; GB:AE00065
A:Experimental source: strain VFS
C:Genetics:
A:Gene: abcT5
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
C:Keywords: ATP; nucleotide binding; P-loop
F:349-543/Domain: ATP-binding cassette homology <ABC>
F:366-373/Region: nucleotide-binding motif A (P-loop)

Query Match 37.2%; Score 1044.5; DB 2; Length 564;
Best Local Similarity 39.7%; Pred. No. 1.5e-66;
Matches 212; Conservative 116; Mismatches 203; Indels 3; Gaps 3;
QY 20 RPAVITAMVPSFINITIALVSPMLQYDRLVLSRVSTLIVTVICVFLFVYGLLEA 79
DB 19 KKTLLIVGLFSLFINILFLPSIYMLAVDLPVSTPFLVITLAVLYFALGLQS 78
QY 80 LRTVILVAGKDFGVADRPFRKSVLDSTLSKRG-IGGAFRMDQVREFMTG-LIAFC 137
DB 79 VRADVMTQISKLDELSEKVEFTSFEKAIKRPKASQPIINDLQKQFLTSVLEFRIE 138
QY 138 DAPRTPVFVIVSMILHPEFGILAIITACIIIFGLAVMNDNATKPIOMATMASIAONDAG 197
DB 139 DLPWPIYFVGLFVHYHYGVMAIISMAVIALILNEYIKKKIKESNELVSTNFLN 198
QY 198 STLNAAEVMKMGWGGOARWRARDEQVAMQAAASDAGAVSGIKVFENIQTLLG 257
DB 199 RALNAAEVEVLDGMRNNLYKRWMPYSKHLISAFEBADRNFLSLTRIFIMOSLMLG 258
QY 258 GGAVLAIIDKISAGAMISILVGRALAPIEGAVGOMKNYIGANGMDRLQTMLEEK-S 316
DB 259 LGVLAIKHEITIGMIVAGSILIGRIDPTIYNGWRQIDNTKAVIRLNEPLKFLFK 318
QY 317 ADHNPRLPERPGVLSAEASILPFGAQQPTMKRASFRIDAGAAVALVSPSAGKSSLLRG 376
DB 319 REVSVKLPERKGEIETLSNVVVVPEEGKTPVLRNINMRILPGEFVALIGPSGSGKSLVRT 378
QY 377 IVGWPCAAGYIRLDGYDIKQWDEKIGRHVGLPQDIELFSGTVNAONIAFTFESEQEV 436
DB 379 ILGIMLPVHGIVEIDGALQKMDRYFGKIVGYLPQDIELFEGVAENIARFGEIDSKTI 438
QY 437 IEAATTLGAVHEMISLPMGYDTAIGEGASISGGRQRLALARAVFRPALLVIDEPNAS 496
DB 439 IEAKLSAHVYIITKLPDGYDTYIGPGITILSGGQRQIALARALYGNPRVYIIDEPDSN 498
QY 497 LDQGEVALMEAMKRLKAKRTVIFATHKVNLAQADYIVINGVISDFGERD 550
DB 499 LDQGEQALYNALIELKKRKYTTIIVSHRIRLNLVDKIAIMODGTLKAFGKAD 552

RESULT 7
A49933
proteinhase SM transporter Prtd - Serratia marcescens
N:Alternate names: inner membrane metalloprotease transporter Prtd
C:Species: Serratia marcescens
C:Date: 13-Sep-1994 #sequence_revision 18-Nov-1994 #text_change 02-Feb-2001
C:Accession: A49933
R:Letoffe, S.; Ghigo, J.M.; Wandersman, C.

J. Bacteriol. 175, 7321-7328, 1993
A:Title: Identification of two components of the Serratia marcescens metalloprotease
A:Reference number: A49933; MUID:94042907
A:Contents: 365
A:Accession: A49933
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-576 <LET>
A:Cross-references: GB:s67013; NID:g452927; PIDN:AAB28772.1; PID:g452928
A:Note: sequence extracted from NCBI backbone (NCBIN:139861, NCBI:P.139863)
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
C:Keywords: ATP; nucleotide binding; P-loop
F:339-523/Domain: ATP-binding cassette homology <ABC>
F:356-363/Region: nucleotide-binding motif A (P-loop)

Query Match 37.0%; Score 1039; DB 2; Length 576;
Best Local Similarity 42.1%; Pred. No. 3.9e-66;
Matches 222; Conservative 96; Mismatches 195; Indels 14; Gaps 5;
QY 28 VESFFINILAVSPMLQYDRLVLSRVSTLIVTVICVFLFVYGLLEALRTQVLR 87
DB 16 LFTAVINILMLAPALYMLQYDRLVPSGNKRTLAMILMAYVGLYFLMGLEWRSQVIR 75
QY 88 GGLKFDGVARDPIKSVLDSTL-SRKIGGQAFRMDQVREFMTG-GLIAPCDAPWTVF 145
DB 76 LGAQMDRLNQRVYDAFETLTKGNPLAGQALNDLNLROFATGNALFAFFDAPFPVY 135
QY 146 VIVSMILHPEFGILAIITACIIIFGLAVMNDNATKPIOMATMASIAONDAG--STRNA 203
DB 136 LTVFLVLPWGLAASAGCDRAGAAGLAQSAQCKRLSAERPSRCANRAGOPANRNA 195
QY 204 EVMRKMGWGGLQARWRARDEQVAMQAAASDAGAVSGIKVFENIQTLLIGGAYLA 263
DB 196 EAIAMGMLDRLRWRARQHQFLQLQNRASEKIAAYTAMSKYTRALQSLMGCGALLA 255
QY 264 IDGKISAGAMISILVGRALAPIEGAVGOMKNYIGARGMDRLQTMLEEKSSADHMLP 323
DB 256 VSGDITFGMIAISILIGRVLGPIDOLIGAMKOWSSAROSLQREVMIAANPRIPSLP 315
QY 324 PEPKGVLSAEASILPFGAQQPTMKRASFRIDAGAAVALVSPSAGKSSLLRGIVGWPC 383
DB 316 PARGALVLSQILTASAPGTAIPVLTGVSFRLEAEVGLVIGASSGTTLLMRQLVGLTP 375
QY 384 AAGYIRLDGYDIKQWDEKIGRHVGLPQDIELFSGTVNAONIAFTFESEQEVIAATLA 443
DB 376 ISG-----DGAQGLPHIGVLPQDIQLPAGTLTDNIARFGEVDARKVVAALALA 425
QY 444 GVHEKIQSLPMGYDTAIGEGASISGGRQRLALARAVFRPALLVIDEPNASIDQGEV 503
DB 426 GVHDLIHLPLPGYETELGEGSGSLSGGQRQVALARALYGSPLAVVIDEPNANDRGEE 485
QY 504 ALMEAMKRLKAKRTVIFATHKVNLAQADYIVINGVISDFGERD 550
DB 486 ALQRIATELAKRGNTIVLVTHKPAIATTDKLVLTAGQOVHGFPSD 532

RESULT 8
T48672
ABC-type transport system ATP-binding protein hasd [validated] - Serratia marcescens
N:Alternate names: hasd export system ABC-type transport component hasd
N:Contains: adenosinetriphosphatase (EC 3.6.1.3)
C:Species: Serratia marcescens
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Sep-2000
C:Accession: T48672
R:Letoffe, S.; Ghigo, J.M.; Wandersman, C.
Proc. Natl. Acad. Sci. U.S.A. 91, 9876-9880, 1994
A:Title: Iron acquisition from heme and hemoglobin by a Serratia marcescens extracellular
A:Reference number: 224524; MUID:95024064
A:Accession: T48672
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-586 <LET>

A:Cross-references: EMBL:X81195; PIDDN:CA57069.1
A:Experimental source: strain S365
C:Genetics:
A:Gene: hasd
C:Complex: the hasa export system consists of the ATP-binding protein hasd (PIR:T48672),
D:P9768101)
C:Function: <ATP>
A:Description: EC 3.6.1.3
C:Function: <HAS>
A:Description: involved in the export of the heme acquisition protein hasa [validated,
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
C:Keywords: Atp binding; hydrolase; membrane-associated protein; protein export

Query Match 37.0%; Score 1039; DB 2; Length 586;
Best Local Similarity 42.1%; Pred. No. 4e-66;
Matches 222; Conservative 96; Mismatches 195; Indels 14; Gaps 5;

OY VESFFINILALVPLMYQVDRLVTSSNSTLVIVYCIEFLVLVGLLEALTQYLVR 87
Db :::||| | ||||| ||||| | | ::::| :| | ||| :|||:|
26 LFTAVINILMLAPALMYQLQVYDRVLPSGNRTIAMLTLMVGLLFEMLLTMSVSQVIR 85
OY GLKFEDGVARDPFEKSYDSFVL-SRKIGIGAFRMDQREPMFG-GLIAECDPAWTFVF 145
Db LGADQMHLNORYDAEFETMKTGNPLAGALNDITLTKRFPAIGNALLFAFDPAWPFPVY 145
OY 146 VIVSMMLHPFEGIIAIIACIIIFGLAVMNDNRKNPDIAOMATSIAMQNDAQ--STLRNA 203
Db 146 LLVVELLHPWLIGALASACDBRAGAAGLAGSACCKRLSEARPGSRCHAFGAOPANLNEA 205
OY 204 EVKAKMGWGGLQARRRRARDQVAWMAASDAGAWSGIKVERNTYTILLGGAYLA 263
Db 206 EAIAIMGWLTJLRIMLMWRHOHQFLLONRASEKIATAVYAMSKTYRLAQSLMICGALLA 265
OY 264 IDGKSACAGMTAGSLTVGRALAPIEGA VGOKNKYTAGARGANDRIOTMRKEKSDDHMP 323
Db 264 :| |:| | ||||||| :| | | :| | | :| | | :| | | :| | | :| | | :| |
266 VSGDTPPEMMTAGSLTIKRVLCAPIDQLLGANKOMSSAQOSIQRLVEVMIANPPRPSP 325
OY 324 PEPRGVLSAEAAISILPPAOOPTMKROASFRI DACAANAVALVPSPAAGKSSLRGIVGPC 383
Db 326 PAPGALLVSQLTASAPCGTAPLYLHGSFRLAGEEVLGVIASGSKTLLMRQLYGALTLP 385
OY 384 AAGVTRLDGYDIKOMDEPKLGRHVGYLPDITEFSGIYAONIARNPTFEEOSEVIETAATLA 443
Db 386 ISG-----DG-----GAEQLPKHIGYLEPODIDLPAQTILTIDNIARFGYDAEKVVAAAAA 435
OY 444 GVHEMIOSLPBGCTPATIGEAGASISGGORORLATARAENFRPALVILDEPARASIDQGEV 503
Db 436 GVHOLITHLPKGYELGESSGLSGCORVALARALAYGSPALVILDEPANLNDREEE 495
OY 504 ALMEARKLKAAKRTVIFATHKVNLQAADVINYINGVISDFGERD 550
Db 496 ALQRATEALKRAGNYIVYTHKRPAIALTTDKLLVLTAGOVGHFEPSP 542

RESULT 9
E98310
rhlzibcJin secretion protein rspd (AF141932) [imported] - Agrobacterium tumefaciens (str
C:Species: Agrobacterium tumefaciens
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C:Accession: E98310
R:Goldman, B., Hinkle, G.; Gatlung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.
A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; PMID:11743194
A:Accession: E98310
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-615 <KUR>
A:Cross-references: GB:AEO07870; PIDN:AAK90007.1; PID:g15159978; GSADB:GN00170
A:Gene: AGR_L_2880

A:Map position: linear chromosome

```
Query Match          36.7%; Score 1029; Db 2; Length 615;
Best Local Similarity 39.2%; Pred. No. 2,2e+5;
Matches 213; Conservative 118; Mismatches 196; Indels 16; Gaps 5;

QY 22 AVITAMVSFF-----INITALVSPLYMLQYDRVLTSRNSTLIVLVICVFLFLV 73
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 17 AVVSPKSTKFFLVALVSGVNNILALTSPFLMQLQYDRLVALAGSLTITGLALLLGLXGF 76
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 74 YGLTALRQTVLVNRGSLKFD---GVARDPIFKSVLSDTLSRKSGGGAFFDMQVREFM 129
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 77 QCLDIIRARVILIRIGEDFDMKRSGRVHDVAVRLPLVNRM--PEDGQLPLRDLONVGF 134
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 130 TG-GLIACDAPWTFVEVIVYSMHLPPFGILAIACIIIFGLAVMNDATNKP1OMATMA 188
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 135 SGTGTAEFFDLPMPLVYLGCLCFLEFHWIGVVALAGAVLVLSITLTITNFSQKPIRDTWVE 194
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 189 STAAONDAGSTLRNNEVKKAGMMGGLQARRRAREDEQVAMQAAASDAGAVMSGIKVFR 248
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 195 NNARRKQLEASRRNAEVVQAMGLGRRLGRKRMQSSNEATVLANRRKAGVAGGICHLASLR 254
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 249 NIVQTLILGGAGYLAIDKISAGAMIAGSLVGRALPIEGAVGQMKNYIGARGAMPRLQ 308
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 255 VYLQSAIILAVGAMVILVREQASGVVMIASSIMMGRALVPLDALISNMKSFVAAPQSMARLR 314
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 309 TMLREKSNADHMLPERRGVLSAFAASILPPGAQQPIMRQASFRIDAGAAVALVGSAA 368
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 315 ELFAQMPANAAAMALPKPEKELRENVITVPGEKKRPVTGLGFVYTAGMLGIIGPSGS 374
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 369 GKSSSLRIGVWPCAAVIRIGDYDIKQMPPEKIRGVGLYLPDIEFSTVQNTARF 428
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 375 GKSTILARILITGMMMAACKVRIDGASFPQMBRETLGRHIGVILPQGVELFPDITGENISRF 434
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 429 TEESQE-VIEAPITAGVHEMIQSLPMGYDTAIGEGASLSGGOROLARAAVFRMPAL 487
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 435 DQNPPEAIIIAAARAAGHEHLIRFEKGYDSDIGAGSTLSAGORIGLARALYGPFI 494
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 488 LVLDPPNALSIOVGVVALMEAMKRLKAKKRVIFATHTVNNLAQADYIMVINOVIDSFG 547
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 495 VVLDPPNMLDAEGEAAVAKAITSVKARGGLAVVVAHRPSAIGAVDYILMEDGRMKAFG 554
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 548 ERD 550
   ||
Db 555 PRD 557

RESULT 10
AF2972
thizlobicin secretion [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AF2972
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo
erage, G.; Gilletto, W.; Grant, C.; Guenther, D.; Kutyavlin, T.; Levy, R.; Li, M.; McCI
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A:Title: The genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AF2972
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-615 <RUB>
A:Cross-references: GB:AE008689; PTDN:AL4196.1; PID:g17741774; GSPDB:GN00187
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: PRD
A:Map position: linear chromosome
```

[illegible][illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2002, 10:42:20 ; Search time 17.51 Seconds

(without alignments)
1227.262 Million cell updates/sec

Title: US-09-913-414-4

Sequence: 1 MFRRSGAKPTFDQAVLVAR.....MNINGVISDPGERDRCPMS 555

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|--------|-------|--------|---------------|---------------------|
| 1 | 1231 | 43.9 | 593 | 1 APRD_PSEAE | Q03024 pseudomonas |
| 2 | 1026.5 | 36.6 | 575 | 1 PRD_ERWCH | P23596 erwinia chr |
| 3 | 584.5 | 20.8 | 707 | 1 HLYB_ACTAC | P23702 actinobacill |
| 4 | 579.5 | 20.6 | 708 | 1 HLYB_PASHA | P16532 pasteurella |
| 5 | 576.5 | 20.5 | 708 | 1 HLYB_PASSP | P55122 pasteurella |
| 6 | 569.5 | 20.3 | 707 | 1 RTIB_ACTPL | P26760 actinobacill |
| 7 | 562.5 | 20.0 | 707 | 1 HLYB_PROVU | P11599 proteus vul |
| 8 | 561.5 | 20.0 | 707 | 1 HLYB_ECOLI | P10089 escherichia |
| 9 | 561.5 | 20.0 | 707 | 1 HLYB_ECOLI | P08716 escherichia |
| 10 | 552.5 | 19.7 | 711 | 1 RT3B_ACTPL | Q04473 actinobacill |
| 11 | 533 | 19.0 | 712 | 1 CVAB_BORPE | P18770 bordetella |
| 12 | 453 | 16.1 | 842 | 1 ABC6_HUMAN | Q9n558 homo sapien |
| 13 | 445.5 | 15.9 | 586 | 1 Y4GB_RHISN | P55469 rhizobium s |
| 14 | 440.5 | 15.7 | 604 | 1 Y4GC_BACSU | P54719 bacillus su |
| 15 | 440 | 15.7 | 582 | 1 MSBA_ECOLI | P27299 escherichia |
| 16 | 419.5 | 14.9 | 587 | 1 MSBA_HAEIN | P44407 haemophilus |
| 17 | 415 | 14.8 | 590 | 1 MDLA_ECOLI | P77265 escherichia |
| 18 | 410 | 14.6 | 722 | 1 MESD_LEUME | Q10418 leuconostoc |
| 19 | 408 | 14.5 | 573 | 1 FTIB_BACSU | P54718 bacillus su |
| 20 | 404 | 14.4 | 735 | 1 ABC8_HUMAN | Q9n418 homo sapien |
| 21 | 401.5 | 14.3 | 694 | 1 ABC7_MOUSE | Q61102 mus musculu |
| 22 | 401.5 | 14.3 | 752 | 1 ABC7_HUMAN | Q75027 homo sapien |
| 23 | 399 | 14.2 | 690 | 1 ATMI_YEAST | P40416 saccharomyc |
| 24 | 394.5 | 14.1 | 631 | 1 YC72_MYCTU | Q11447 mycobacteri |
| 25 | 394.5 | 14.1 | 1321 | 1 MDR1_CAEEL | P34712 caenorhabdi |
| 26 | 392.5 | 14.0 | 715 | 1 LCCL_LACLA | Q9CJb8 lactococcus |
| 27 | 391.5 | 13.9 | 598 | 1 Y288_THEMEA | Q9WY44 thermotoga |
| 28 | 390.5 | 13.9 | 698 | 1 CVAB_ECOLI | P25220 escherichia |
| 29 | 388.5 | 13.8 | 859 | 1 YDA8_MYCTU | Q11018 mycobacteri |
| 30 | 386 | 13.8 | 583 | 1 EXP8_STRPN | P35598 streptococc |
| 31 | 386 | 13.8 | 717 | 1 COMA_STRPN | Q03727 streptococc |
| 32 | 380.5 | 13.6 | 1279 | 1 MDR3_HUMAN | P21339 homo sapien |
| 33 | 380.5 | 13.6 | 1905 | 1 TAGB_DICDI | P54683 dictyostell |

| | | | | | |
|----|-------|------|------|--------------|--------------------|
| 34 | 379.5 | 13.5 | 1254 | 1 MDR3_CAEEL | P34713 caenorhabdi |
| 35 | 379.5 | 13.5 | 1280 | 1 MDR1_HUMAN | P08183 homo sapien |
| 36 | 378.5 | 13.5 | 715 | 1 LCNC_LACLA | Q00564 lactococcus |
| 37 | 377 | 13.4 | 693 | 1 YET1_SCHPO | Q14286 schizosacch |
| 38 | 376.5 | 13.4 | 1276 | 1 MDR2_CRIGR | P21449 cricetus |
| 39 | 375.5 | 13.4 | 575 | 1 YWJA_BACSU | P45861 bacillus su |
| 40 | 374.5 | 13.3 | 1281 | 1 MDR3_CRIGR | P23174 cricetus |
| 41 | 369.5 | 13.2 | 1743 | 1 TAGC_DICDI | Q2368 dictyostell |
| 42 | 367.5 | 13.1 | 1276 | 1 MDR1_CRIGR | P21448 cricetus |
| 43 | 366.5 | 13.1 | 1276 | 1 MDR2_MOUSE | P21440 mus musculu |
| 44 | 366.5 | 13.1 | 1278 | 1 MDR2_RAT | Q08201 rattus norv |
| 45 | 365.5 | 13.0 | 584 | 1 LMRA_LACLA | Q9ch18 lactococcus |

ALIGNMENTS

| RESULT | 1 | STANDARD | PRT | 593 AA. |
|-----------------|--|----------|-----|---------|
| APRD_PSEAE | | | | |
| ID | APRD_PSEAE | | | |
| AC | Q03024; | | | |
| DT | 01-OCT-1993 (Rel. 27, Created) | | | |
| DT | 01-OCT-1993 (Rel. 27, Last sequence update) | | | |
| DT | 16-OCT-2001 (Rel. 40, Last annotation update) | | | |
| DE | Alkaline protease secretion ATP-binding protein aprd. | | | |
| GN | APRD OR PA1246. | | | |
| OS | Pseudomonas aeruginosa. | | | |
| OC | Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae; | | | |
| OX | Pseudomonas. | | | |
| NCBI_TaxID=287; | | | | |
| [1] | | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-ATCC 15692 / PA01; | | | |
| RX | MEDLINE-93051361; Pubmed-1427098; | | | |
| RA | Duong F., Lazdunski A., Caml B., Murgier M.; | | | |
| RT | "Sequence of a cluster of genes controlling synthesis and secretion | | | |
| RT | of alkaline protease in Pseudomonas aeruginosa: relationships to | | | |
| RT | other secretory pathways."; | | | |
| RL | Gene 121:47-54(1992). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-ATCC 15692 / PA01; | | | |
| RX | MEDLINE-20437337; Pubmed-10984043; | | | |
| RA | Strover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P., | | | |
| RA | Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M., | | | |
| RA | Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y., | | | |
| RA | Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M., | | | |
| RA | Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T., | | | |
| RA | Reitzer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.; | | | |
| RT | "Complete genome sequence of Pseudomonas aeruginosa PA01, an | | | |
| RT | opportunistic pathogen."; | | | |
| RL | Nature 406:959-964(2000). | | | |
| CC | -1- FUNCTION: INVOLVED IN THE SECRETION OF ALKALINE PROTEASE. | | | |
| CC | -1- SUBCELLULAR LOCATION: Integral membrane protein. | | | |
| CC | -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. HLYB SUBFAMILY. | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ | | | |
| CC | or send an email to license@isb-sib.ch). | | | |
| CC | ----- | | | |
| DR | EMBL; X64558; CAA45855.1; - | | | |
| DR | EMBL; AE004554; AAG04635.1; - | | | |
| DR | PIR; S26696; S26696. | | | |
| DR | HSSP; P13569; INBD. | | | |
| DR | InterPro: IPR003593; AAA. | | | |
| DR | InterPro: IPR001140; ABC_transporter_tmem. | | | |
| DR | InterPro: IPR003439; ABC_transporter. | | | |
| DR | InterPro: IPR001687; ATP_GTP_A. | | | |
| DR | Pfam; PF00664; ABC_membrane; 1. | | | |

| ID | HLVB_ACTAC | STANDARD: | PRT: | 707 AA. |
|----|---|-----------|------|---------|
| AC | HLVB_ACTAC | | | |
| DT | P23702; | | | |
| DT | 01-NOV-1991 (Rel. 20, Created) | | | |
| DT | 01-NOV-1991 (Rel. 20, Last sequence update) | | | |
| DT | 16-OCT-2001 (Rel. 40, Last annotation update) | | | |
| DE | Leukotoxin secretion ATP-binding protein. | | | |
| GN | LKTB OR LTB. | | | |
| OS | Actinobacillus actinomycetemcomitans (Haemophilus actinomycetemcomitans). | | | |
| OC | Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae; | | | |
| OC | Actinobacillus. | | | |
| OX | NCBI_TaxID=714; | | | |
| RA | [1] | | | |
| RA | SEQUENCE FROM N.A. | | | |
| RA | MEDLINE=90384842; PubMed=2402457; | | | |
| RA | Guthmiller J.M., Kolodrubetz D., Cagle M.P., Kraig E.; | | | |
| RT | "Sequence of the lktB gene from Actinobacillus actinomycetemcomitans." | | | |
| RL | Nucleic Acids Res. 18:5291-5291(1990). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RP | MEDLINE=92072446; PubMed=1961107; | | | |
| RA | Lally E.T., Golub E.E., Kleba I.R., Taichman N.S., Decker S., | | | |
| RA | Berthold P., Gibson C.W., Demuth D.R., Rosenblom J.; | | | |
| RT | "Structure and function of the B and D genes of the Actinobacillus actinomycetemcomitans leukotoxin complex." | | | |
| RL | Microb. Pathog. 11:111-121(1991). | | | |
| CC | -1- FUNCTION: INVOLVED IN THE EXPORT OF LEUKOTOXIN (HEMOLYSIN). | | | |
| CC | -1- SUBCELLULAR LOCATION: Integral membrane protein. | | | |
| CC | -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. HLVB SUBFAMILY. | | | |
| CC | | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/). | | | |
| CC | or send an email to license@isb-sib.ch . | | | |
| CC | | | | |
| DR | EMBL; X53955; CAA37906.1; -. | | | |
| DR | PIR; S12601; S12601. | | | |
| DR | PIR; A61378; A61378. | | | |
| DR | InterPro: IPR003593; AAA. | | | |
| DR | InterPro: IPR001140; ABC transporter_tmam. | | | |
| DR | InterPro: IPR001439; ABC transporter. | | | |
| DR | InterPro: IPR001687; ATP_GTP_A. | | | |
| DR | Pfam; PF00664; ABC_membrane; 1. | | | |
| DR | Pfam; PF00005; ABC_tran; 1. | | | |
| DR | SMART; SMO0382; AAA; 1. | | | |
| DR | PROSITE; PS00211; ABC_TRANSPORTER; 1. | | | |
| KW | Hemolysis; Transport; ATP-binding; Transmembrane. | | | |
| PT | NP_BIND 502 509 ATP (BY SIMILARITY). | | | |
| SO | SEQUENCE 707 AA: 79578 MW: 75564 ECAC46DF4B CRC64: | | | |

| | | | | | | |
|----|-----------------------|--|------------------|------------|------------|--|
| | Query Match | 20.8% | Score 584.5 | DB 1 | Length 707 | |
| | Nest Local Similarity | 30.0% | Pred. NO.5.4e-33 | | | |
| | Matches 169, | Conservative 106; | Mismatches 247; | Indels 41; | Gaps 12. | |
| Oy | 10 | TTFDQAVLVARPAVITAMVFSPFINILALVSPLMQYDRVLRNWSLTIVLVICVF | 69 | | | |
| | | : : : : : : : : | | | | |
| Dd | 141 | TWFIIPAVIKRYKRIFETFLYSIFLDIIPALIPPLPFQYVMKVLVRGSESTLVIIIVALLAI | 200 | | | |
| | | | | | | |
| Oy | 70 | LFLVYGLLEALRQTVLVARGGLKFEDGARDPIFKSVLDLSL-----RKIGSQAFRRMDQ | 124 | | | |
| | | : : : : : : : : : : : : : : : : : | | | | |
| Dd | 201 | VLFEEILLGGRTTYFAHSTSRIDVELGARLERRHLLAIPISFEARRRGDTVARVERDQ | 260 | | | |
| | | | | | | |
| Oy | 125 | VREFTTG-GLIAFCADPMTVPFVIYSMMMLHP--FFGIILAIICILIEFGLAV-----MN | 174 | | | |
| | | : : : : : : : : : : : : : : : : : : | | | | |
| Dd | 261 | IRNFILITGQLNTSIDLLRSFIFFANWMYVSRLTIVLGSLPCYIYWTSFIIPIRRRD | 320 | | | |
| | | | | | | |
| Oy | 175 | DNAIKN-PIQMATMSIAQNDDAGSTLRNAAEYVKRAMGMINGLIQAQRRAARDEQVAMAQA | 233 | | | |
| | | | | | | |

```

Db      321 DKFARNADNOSFVLESTYAIN-----TIAAMISFOMNINW---DKOLASTYAV 366
Qy      234 SDAGGAV-----MSGITKFRNIVOTLLILGGAGYLAIDKISAGAMIASILVGRALPIEG 289
Db      367 SFKYTVLATIQOOGIQLQKRVAMVINMLGHLVYISGDISIGOLLAFNMLAGQISVIR 426
Qy      290 AVGQWKNYIGARGAMDRLQTMLR---EKSADMDPLPREBRYLSAEFASILPPGAOQPT 346
Db      427 LAQIMQDPQYQGISVTRIGDVLSNPENTENNTAS--VSLPEYQELGSFENIKFRKKPPDSPMI 484
Qy      347 MKQASFRIDAGAAVAALVPSASGKSSLLRGIVGWPACAGVIRLDGYDIKQMPDKLGRH 406
Db      485 LNNIMLDSIGEVIGIVGSRSGSKSTLTKLQRRYIPREQGVLLDGHDLADPNMLRQ 544
Qy      407 VGYLPQDIELESGTVNAONIAFTE-FFSQEYIEAATLAGVHEMTOISLPMGYDTALGEGGA 465
Db      545 VGVVLQDVNLNRSIRIENIALTNPNGMPEMEKVIAAKLAGAHDFISELRGEGYNVWGEGA 604
Qy      466 SLSGGQRRRLAALAVFMRPLVLDVERPNMSLDVOGVGVALMEAKKRLAAKRYIVFTTK 525
Db      605 GLSGGQRRRIALAVLNNPRILFLDEATSLADYENSIIMNNHKI-CONRTVLLIATHR 663
Qy      526 VNLLAADYIMVIVGSDGFE 548
Db      664 LSTVKNADRIITVMDKGETIEGK 686

```

| | | | | |
|------------|---|-----------|------|---------|
| RESULT | 4 | | | |
| HLVY_PASHA | | | | |
| ID | HLVY_PASHA | STANDARD; | PRT; | 708 AA. |
| AC | P16532. | | | |
| DT | 01-AUG-1990 (Rel. 15, Created) | | | |
| DT | 01-AUG-1990 (Rel. 15, Last sequence update) | | | |
| DT | 16-OCT-2001 (Rel. 40, Last annotation update) | | | |
| DE | Leukotoxin secretion ATP-binding protein. | | | |
| GN | LkTx. | | | |
| OS | Pasteurella haemolytica. | | | |
| OC | Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae; | | | |
| OC | Mannheimia. | | | |
| OX | NCBI_Taxid=75985; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-SEROTYPE AI. | | | |
| RX | MEDLINE=87306837; Pubmed=3040588; | | | |
| RA | Lo R.Y.C., Strathdee C.A., Shewen P.E.; | | | |
| RT | "Nucleotide sequence of the leukotoxin genes of Pasteurella | | | |
| RT | haemolytica AI."; | | | |
| RL | Infect. Immun. 55:1987-1996(1987). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-SEROTYPE AI. | | | |
| RX | MEDLINE=89123172; Pubmed=2914876; | | | |
| RA | Strathdee C.A., Lo R.Y.C.; | | | |
| RT | "Cloning, nucleotide sequence, and characterization of genes encoding | | | |
| RT | the secretion function of the pasteurella haemolytica leukotoxin | | | |
| RT | determinant."; | | | |
| RL | J. Bacteriol. 171:916-928(1989). | | | |
| RN | [3] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-SEROTYPE AI / PHL101; | | | |
| RX | MEDLINE=89210283; Pubmed=2707120; | | | |
| RA | Highlander S.K., Childamaram M., Engler M.J., Weinstein G.M.; | | | |
| RT | "DNA sequence of the Pasteurella haemolytica leukotoxin gene | | | |
| RT | cluster."; | | | |
| RL | DNA 8:15-28(1989). | | | |
| RN | [4] | | | |
| RP | SEQUENCE OF 1-48 FROM N.A. | | | |
| RX | MEDLINE=90236888; Pubmed=2185213; | | | |
| RA | Highlander S.K., Engler M.J., Weinstein G.M.; | | | |
| RT | "Secretion and expression of the Pasteurella haemolytica Leukotoxin." | | | |
| RL | J. Bacteriol. 172:3343-3350(1990). | | | |
| CC | -1- FUNCTION: INVOLVED IN THE EXPORT OF LEUKOTOXIN. | | | |

```
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. HLYB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
```

DR EMBL; M20730; AAA25530.1; -;
DR EMBL; M24197; AAA25544.1; -;
DR PIR; A32051; A32051.
DR PIR; S29517; S29517.
DR InterPro; IPR003593; AAA.
DR InterPro; IPR001140; ABC_transporter_tmem.
DR InterPro; IPR003439; ABC_transportr.
DR InterPro; IPR001687; ATP_GTP_A.
DR Pfam; PF00664; ABC_membrane; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW Hemolysis; Transport; ATP-binding; Transmembrane.
FT TRANSMEM 216 236 POTENTIAL.
FT TRANSMEM 249 269 POTENTIAL.
FT TRANSMEM 327 347 POTENTIAL.
FT TRANSMEM 353 373 POTENTIAL.
FT TRANSMEM 446 466 POTENTIAL.
FT TRANSMEM 468 488 POTENTIAL.
FT NP_BIND 503 510 ATP (BY SIMILARITY).
FT NP_BIND 512 512 Q -> T (IN REF. 3).
FT CONFLICT 340 340 S -> A (IN REF. 3).
FT CONFLICT 366 367 SS -> FI (IN REF. 3).
FT CONFLICT 597 597 N -> T (IN REF. 3).
SQ SEQUENCE 708 AA; 79712 MW; BE2ABAF12380A4E CRC64;

Query Match 20.6%; Score 579.5; DB 1; Length 708;
Best Local Similarity 29.6%; Pred. No. 1.2e-32;
Matches 166; Conservative 116; Mismatches 243; Indels 35; Gaps 14;

QY 10 TTFDQAVLVARPAVTAMVSEFFINITALVSPLYMLQYDRVITSNNSTLVLYVICVF 69
DB 142 TWEIFAVIKYRKIFETLIVSIFLQIFALITPLEFOVYMDKVLVHGFSTLNIITVALAI 201
QY 70 LFLVYGLLEALRTOYLVNRGLKFGDVGARDPIKSVLDSTLS-----RKIGGAQAFRDMQ 124
DB 202 VLIETIVISGLRTYVFSHSTSRIDVELGAKLFRHLISLPISFENRRVGDIVARRELDQ 261
QY 125 VEEFMTG-GLIAFCDAPTVPFVIVSWMLHP--FFGIIAIIACIIIFGLAVNNDATKNP 181
DB 262 IKNFLTGALTSVLDLTFEIFFAWMYSPKLTIVLTIGSLCYILMSIFI-----SP 314
QY 182 I---QMATVASTAONDA--GSTLRNAEYKAKMGGLQARRRARRDQVAAQASPA 236
DB 315 IIRRLDEKFAASADNOAPLVESVTSINNIKAMAAVAPQMTDW---DKQLASVSSSPR 370
QY 237 GGAIV-----MSGIKVFNPIYQITLILGGAYLAIDKISAGAMAGSILVGRALPIEGANG 292
DB 371 VYVLATIGGQGVOLQKQTVAVNLMLGAHLVIGDLSIQLLAFNLSQVLAIPRLAQ 430
QY 293 QMKNYTGARGADRLQTMLE- EKASADHMLPEPRGVLSAEAAST-LPPGAQOPT-MKQ 349
DB 431 LMDFQGVGISTVRLGDLVNSPFEQYQKLSLPEIKGDISFKNIIRRYKRPDA--PILNN 488
QY 350 ASFRIDAGAAVALVGPASAKSSLLRGIVGWPSCAGVIRLDYDIKQMDPEKLGHVGY 409
DB 489 VMLIEIGQEVIGIVSGSGSKSTLKLORFYIPENGQVLAIDHDALADPMMLRQIOIGV 548
QY 410 LQDIEFLFGTYAQNARTLE-FESQEVLEAATLAVHMIQSLPBGYDTAIGEGASIS 468
DB 549 VLDNVLNRSIRENTALSDPGMPMERVIYAALAGAHDFISELRGQNTIVIGEAGGIS 608

QY 469 GGGORIALARAFRMPALVDEPNASLDQGVVALMEAMKRLAKRTVIFATHKVNL 528
DB 609 GGGORIALARAFRMPALVDEPNASLDQGVVALMEAMKRLAKRTVIFATHKVNL 528
QY 529 LAQDYIMVINOQVTSDFGE 548
DB 668 VKNADRIIVMEKGEIVGOK 687

RESULT 5
HLYB_PASSP STANDARD; PRT; 708 AA.
ID HLYB_PASSP STANDARD; PRT; 708 AA.
AC P55122;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Leukotoxin secretion ATP-binding protein.
GN LKTB.
OS Pasteurella haemolytica-like sp. (strain 5943B).
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=28165;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93239320; PubMed=8478098;
RA Chang Y.-F., Ma D.-P., Shi J., Chengappa M.M.;
RT "Molecular characterization of a leukotoxin gene from a Pasteurella
RT haemolytica-like organism, encoding a new member of the RTX toxin
RT family.";
RL Infect. Immun. 61:2089-2095(1993).
RL -1- FUNCTION: INVOLVED IN THE EXPORT OF LEUKOTOXIN.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. HLYB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC EMBL; L12148; AAA16445.1; -;
DR InterPro; IPR003593; AAA.
DR InterPro; IPR001140; ABC_transporter_tmem.
DR InterPro; IPR003439; ABC_transportr.
DR InterPro; IPR001687; ATP_GTP_A.
DR Pfam; PF00664; ABC_membrane; 1.
DR Pfam; PF00005; ABC_tran; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW Hemolysis; Transport; ATP-binding; Transmembrane.
FT TRANSMEM 159 176 POTENTIAL.
FT TRANSMEM 189 212 POTENTIAL.
FT TRANSMEM 270 288 POTENTIAL.
FT TRANSMEM 295 316 POTENTIAL.
FT TRANSMEM 388 404 POTENTIAL.
FT TRANSMEM 411 428 POTENTIAL.
FT NP_BIND 503 510 ATP (BY SIMILARITY).
SQ SEQUENCE 708 AA; 79772 MW; 7326D21411B5090A CRC64;

Query Match 20.5%; Score 576.5; DB 1; Length 708;
Best Local Similarity 29.2%; Pred. No. 1.9e-32;
Matches 165; Conservative 116; Mismatches 243; Indels 41; Gaps 14;

QY 10 TTFDQAVLVARPAVTAMVSEFFINITALVSPLYMLQYDRVITSNNSTLVLYVICVF 69
DB 142 TWEIFAVIKYRKIFETLIVSIFLQIFALITPLEFOVYMDKVLVHGFSTLNIITVALAI 201
QY 70 LFLVYGLLEALRTOYLVNRGLKFGDVGARDPIKSVLDSTLS-----RKIGGAQAFRDMQ 124

```
Db 202 VIREIYLSGLRTYIFAHSTSRIDVELGARLFRLHLALPISYFENRRVGDVAVARELDQ 261
Qy 125 VREMTG-GLIAFCDAWPTPEVIVSMNLHP--FGILAIACIIINGLAN-----MN 174
Db 262 IRNFLTGQALTSVLDLMSFFFAVMWYSPKLTVLVIGSLPCYILMSIFSPILRRRLD 321
Qy 175 DNATKNTIOMATMA-SIAQNDAGSTLRNAEVMKAMGMGGLQARWRARDEQVAMQAAA 233
Db 322 DKFRAGDNQALFVSEVTAIN-----MIKAMVSPQMTWTW---DKQIASYVSS 367
Qy 234 SDAGGAV---MSGIKVFRNIYOTLLIGGAYLAIDKISAGAMIAGSILVGRALAPLEG 289
Db 368 SFRVTYVATIGQGVQVLIQKTVWYINLMGLAHVLISGDLISGILAFRNMLSGVYIAPYIR 427
Qy 290 AVGOMKVIYARGAMDRLOQMLRE-EKSADDMPLPEPRGVLSAEASI-LPPAQOQPT- 346
Db 428 LAOLMDFOTGVGISVTLGDVLSPTBOYQCKLSLPETQGDIAKNIREFRYKPPDA--PTI 485
Qy 347 MRQASFRIDGAAVALVGPSACKSSLRGIVGWPCAGVIRLDGYDKOMDEKLGRIH 406
Db 486 LNNVNLTKKGEVIGIVRGSSGKSTLTLLQRYIPENGVGLDGHDLALADNMLRRQ 545
Qy 407 VGIYLPDIELEPGTVAONIAFTEFESOE-VIEATLAGVHEMTQSLPMGYDAIGEGGA 465
Db 546 IGVVLQDNVLNRSIRENIALSEPGMSMERIYAAKLGAHDFISDVEGNYITVIEGGA 605
Qy 466 SLGGGQRLARAVFRMALLVLDENPNSLDQVGEVALMEAKRKAKRTYIFATHK 525
Db 606 GLSSGGQRIARALAVLNPNRLLFEDRTSALDSESHIIMQNMOKI-CQGRVILIAHR 664
Qy 526 VNLLAQADYIMVINOYISDFGERD 550
Db 665 LSTYKNADRIIVMEKGEIVGGRKN 689

RESULT 6
RTLB_ACRPL STANDARD; PRT; 707 AA.
ID RTLB_ACRPL
AC P26760;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE RTX-I toxin determinant B (Toxin RTX-I secretion ATP-binding protein)
DE (APX-IB) (CLY-IB) (CYCLOYSIN IB) (CLY-IB).
GN APXIB OR CLYIB OR HLYIB OR APPB
OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellales;
OC Actinobacillus.
OX NCBI_TaxID=715;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SEROTYPE 5;
RX MEDLINE-91317735; PubMed-1860823;
RA Chang Y.-F., Young R., Struck D.K.;
RT The Actinobacillus pleuropneumoniae hemolysin determinant: unlinked
RT appca and appbd loci flanked by pseudogenes.";
RL J. Bacteriol. 173:5151-5158(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ISOLATE CVI 13361 / SEROTYPE 9;
RX MEDLINE-92040145; PubMed-1937809;
RA Smits M.A., Briatore J., Jansen R., Smith H.E., Kamp E.M.,
RA Gielkens A.L.J.;
RT Oxytolysins of Actinobacillus pleuropneumoniae serotype 9.";
RL Infect. Immun. 59:4497-4504(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-S 4074 / SEROTYPE 1;
RX MEDLINE-94237497; PubMed-8181764;
RA Frey J., Haldemann A., Nicolet J., Boffin A., Prentki P.;
RT "Sequence analysis and transcription of the apxi operon (hemolysin I)
RT from Actinobacillus pleuropneumoniae.";
```

```
RL Gene 142:97-102(1994).
CC -1- FUNCTION: INVOLVED IN THE TRANSPORT OF THE TOXIN RTX-I AS WELL AS
CC THAT OF RTX-II.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF SEROTYPE 9.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. HLYB SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M65808; AAB00966.1; -
DR EMBL; X61112; CAA43425.1; -
DR EMBL; X68595; CAA48587.1; -
DR PIR; A40366; A40366.
DR PIR; S18855; S18855.
DR InterPro; IPR003593; AAA.
DR InterPro; IPR001140; ABC_Transporter_tmem.
DR InterPro; IPR003439; ABC_transportr.
DR InterPro; IPR001687; ATP_GTP_A.
DR Pfam; PF00664; ABC_membrane.1.
DR Pfam; PF00005; ABC_tran.1.
DR SMART; SM00382; AAA.1.
DR PROSITE; PS00211; ABC_TRANSPORTER.1.
DR Hemolysis; Cytolysis; Transport; ATP-binding; Transmembrane.
KW TRANSMEM 158 179
FT TRANSMEM 187 204
FT TRANSMEM 269 289
FT TRANSMEM 293 311
FT TRANSMEM 387 403
FT NP_BIND 502 509
FT VARIANT 110 112 EOA -> KQT (IN SEROTYPE 5).
FT VARIANT 328 328 D -> H (IN SEROTYPE 5).
FT VARIANT 517 517 R -> V (IN SEROTYPE 5).
FT VARIANT 556 556 N -> G (IN SEROTYPE 5).
FT VARIANT 609 611 GQR -> RAT (IN SEROTYPE 1).
FT VARIANT 611 612 DR -> PN (IN SEROTYPE 5).
FT VARIANT 671 672 DR -> AS (IN SEROTYPE 5).
FT VARIANT 692 692 A -> R (IN SEROTYPE 5).
SQ SEQUENCE 707 AA; 79663 MW; 34A4339C57340DE9 CRC64;

Query Match 20.3%; Score 569.5; DB 1; Length 707;
Best Local Similarity 28.4%; Pred. No. 5.8e-32;
Matches 161; Conservative 105; Mismatches 253; Indels 47; Gaps 9;

Qy 10 TIFQAVLVARPAVITAMVFSFINILALVSPMLQYDVNLSRVNSTIVLTVCVF 69
Db 141 TWPIPAVYKRIKIFETLIVSIFLQIFALTPLEFYQVMDLVHRGSELTIVTVALAI 200
Qy 70 LFLVYGLLEALRTQVLRVGGKFGVARDPIFKSLVDSTLS-----RKGIQGAFRWDQ 124
Db 201 VLEIYVINGLRTYIFAHSTSRIDVELGARLFRLHLALPISYFENRRVGDVAVARELDQ 260
Qy 125 VREMTG-GLIAFCDAWPTPEVIVSMNLHPFGIILAIACIIIFGLAVMNDNAFKNPID 183
Db 261 IRNFLTGQALTSVLDLMSFFFAVMWYSPKLTVLVIGSLPCYILMSIFSPILRRRLD 320
Qy 184 MATMASTAAQNDAGSTLRNAEVMKAMGMGGLQARWRARDEQVAMQAAA-----SDAGGA 239
Db 321 EKFRAGDNQSELVSEVTAINITIKALAVTPQMTWTW---DKQIASYVSAEFRVYTLATI 376
Qy 240 VMSGIKVRNIYOTLLIGGAYLAIDKISAGAMIAGSILVGRALAPLEGAVGOMKNVIG 299
Db 377 GQGVQVLIQKTVWYINLMGLAHVLISGDLISGILAFRNMLSGVYIAPYIRLAOLMDFQ 436
Qy 300 ARGAMDRLOQMLRE-EKSADDMPLPEPRGVLSAEASII-LPPAQOQPTMRQASFR----- 353
```

```

Db 437 VGISVTRGLDVNSPTESYOGKLAPEIKGDI-----TFRNIRFRKPKDA 481
QY 354 -----IDGAAVALVGPSSAKSSLLRIGVWPCAGVIRLUDGDIKOMDEKL 403
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 482 PVLNDVNLSTIOGEEVIGVGRSGSKSTILKIQRYIPENGOVLIDGDLADPNWL 541
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 404 GRHVGLYLPDIEFSGTVAONIAFTE--FESQEVIEATLAGVHEMIQSLPMGDTAIGE 462
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 542 RROGVVYLQDNVLLNRSIRDNIALADGPMPEKIVHAAKLAGAHEFISELREGYNTIGE 601
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 463 GGASLSGGQORALAAVFRMALLVLDEPNMSLDQVGEVALMEAKRKLAARTYIFA 522
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 602 QGAGLSGGQORALAAVFRMALLVLDEPNMSLDQVGEVALMEAKRKLAARTYIFA 522
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 523 THKVNLLAODYIMVINOQVYISDGE 548
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 661 AHRISTYKNNADRIIVMEKGQIVEGK 686

RESULT 7
HLV2_PROVU STANDARD; PRT; 707 AA.
AC P11599;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hemolysin secretion Amp-binding protein.
GN HLYB.
OS Proteus vulgaris.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Proteus.
OX NCBI_TaxID=585;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9039746; Pubmed=3054490.
RA Kotonakis V., Kotonakis E., Hughes C.;
RT "Comparison of the haemolysin secretion protein HlyB from Proteus
RT vulgaris and Escherichia coli; site-directed mutagenesis causing
RT impairment of export function.";
RL Mol. Genet. 213:551-555(1988).
CC -1- FUNCTION: INVOLVED IN THE EXPORT OF HEMOLYSIN A.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. HLYB SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X12852; CA31330.1; -.
DR PIR: S03477; LEEBBV.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR001140; ABC_transporter_tmnm.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam: PF00664; ABC_membrane; 1.
DR SMART: SM00005; ABC_tran; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
KW Hemolysis; Transport; ATP-binding; Transmembrane.
FT TRANSMEM 158 179 POTENTIAL.
FT TRANSMEM 187 204 POTENTIAL.
FT TRANSMEM 269 289 POTENTIAL.
FT TRANSMEM 293 311 POTENTIAL.
FT TRANSMEM 360 377 POTENTIAL.
FT TRANSMEM 387 403 POTENTIAL.
FT TRANSMEM 407 425 POTENTIAL.
FT NP_BIND 502 509 ATP (BY SIMILARITY).
SQ SEQUENCE 707 AA; 79940 MW; 6972C3D4EB18204F CRC64;

```

```

Query Match 20.0%; Score 562.5; DB 1; Length 707;
Best Local Similarity 29.7%; Pred. No. 1,8e-31;
Matches 170; Conservative 102; Mismatches 240; Indels 61; Gaps 14;

QY 10 TIEDQAVLVRPAVITAMVSPFINIALVSPLYMQVYRVLSRVNSTLYIVYICVF 69
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 141 TWETPAVKRYKRIETELIVSVFLQLELTPLEFOVVMKVLVHRGSESTLITITLAV 200
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 70 LEIVYGLLEALRTQVILRGGLKEGVARDPIFKSVLDSTLS-----RKIGGOAFRMDQ 124
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 201 VAIFETILSGERTYIFTHSRIDVELGAKLFRLILALPISTYESRRVGDTVARVELDQ 260
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 125 VREFMTG-GLIACFDAPWTPVEVYSWMLHP--FEGILAIACILIFGLAV-----MN 174
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 261 IRNELTQALTSIIDLFSEFFVAVMYSPKLTLYILFSLPCATWISFISPLRRRLD 320
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 175 DNATKN-PIQMATMASIAQNDAGSTLRNNEVMKMGWGGLQARVARRDEQVAMQAAA 233
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 321 DKFARNADNOSFLVESVTAIN-----TKAMAVSPQMTNIV---DKQLAGTYA- 365
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 234 SDAGGAV-----MSGIKYFRNIVOTLLIGGAYLAIDKISAGAMIAGSIIVGRALAP 286
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 366 --AGFKTVLATIGQGGIQLQKRAVMINIMLGHVITSGLISGULAFNMLAGQIVAP 423
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 287 IEGAVGQWKNYIGARGAMDRLQMLR--EKSADDMPLPEPRGVL-----AQAAS 336
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 424 VIRLAQLMODFOQVIGSVTRGLDVNSPTESYOGKLTLPETINGDISPRNIRFRKYPAPPI 483
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 337 ILPPGAQOPTRKQAFRIIDGAAVALVGPSSAKSSLLRIGVWPCAGVIRLUDGDIK 396
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 484 IL-----NNINLNKQGEITIGVGRSGSKSTILKIQRYIPENGOVLIDGDLA 534
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 397 QMDEPKLRHVGLYLPDIEFSGTVAONIAFTE--FESQEVIEATLAGVHEMIQSLPMG 455
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 535 LADNNMLRQGVVYLQDNVLLNRSIRDNIALADGPMPEKIVHAAKLAGAHEFISELREG 594
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 456 YDTAIGEGASLSGGQORALAAVFRMALLVLDEPNMSLDQVGEVALMEAKRKLA 515
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 595 YNTIVGEGAGLSGGQORALAAVFRMALLVLDEPNMSLDQVGEVALMEAKRKLA-CQ 653
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 516 KRYVIFATHKVNLLAODYIMVINOQVYISDGE 548
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 654 GRVYIIIAHRLSTYKNNADRIIVMEKGQIVEGK 686

RESULT 8
HLV2_ECOLI STANDARD; PRT; 707 AA.
AC P10089;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hemolysin secretion ATP-binding protein, chromosomal.
GN HLYB.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=J96 / SEROTYPE O4;
RX MEDLINE=85234404; Pubmed=3891743;
RA Felmele T., Pellett S., Welch R.A.;
RT "Nucleotide sequence of an Escherichia coli chromosomal hemolysin.";
RL J. Bacteriol. 163:94-105(1985).
CC -1- FUNCTION: INVOLVED IN THE EXPORT OF HEMOLYSIN A.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. HLYB SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```


the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; M10133; AAA23976.1; -
 CC PIR; B24433; LEECB.
 DR InterPro: IPR003593; AAA.
 DR InterPro: IPR001140; ABC_transporter_tmemb.
 DR InterPro: IPR003439; ABC_transporter.
 DR InterPro: IPR001687; ATP_GTP_A.
 DR Pfam; PF00664; ABC_membrane; 1.
 DR Pfam; PF00005; ABC_tran; 1.
 DR SMART; SMO0382; AAA; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 KM Hemolysis; Transport; ATP-binding; Transmembrane.
 FT TRANSMEM 158 179. POTENTIAL.
 FT TRANSMEM 187 204. POTENTIAL.
 FT TRANSMEM 269 289. POTENTIAL.
 FT TRANSMEM 293 311. POTENTIAL.
 FT TRANSMEM 360 377. POTENTIAL.
 FT TRANSMEM 387 403. POTENTIAL.
 FT TRANSMEM 407 425. POTENTIAL.
 FT NP_BIND 502 509. ATP (BY SIMILARITY).
 SQ SEQUENCE 707 AA; 79463 MW; 21099CB45E59437E CRC64;

Query Match 20.0%; Score 561.5; DB 1; Length 707;
 Best Local Similarity 29.7%; Pred. No. 2.1e-31;
 Matches 172; Conservative 101; Mismatches 233; Indels 73; Gaps 14;

10 TIFQAVLVAPAVITANVFSEFINIILALVLYDLYSRNVTSLVILVICVF 69
 141 TWFPALIKYKITEVLIVSVFQLFLITPLFFQVMDKVLVHGRSTLAVITVALSV 200
 70 LFLVYGLLEALRTQVLVNGKLPDGVARDPLFKSVLSTLSL-----RKIGGQAFRDMQ 124
 201 VVVEFIILSGLRJYIFANSTSRIDVELGAKLFRLHLALPIVFESERRVGDVAVARELDQ 260
 125 VREMTG-GLAFCDAFTPVFVYSMLHP--FFGILAIACIIIFGLAV-----MN 174
 261 IRNPLTGALTSVLDLFSFIFFAVMYSPKLTLYLIFSPCVAAWGVFTSPILRRILD 320
 175 DNATKN-PIDWATNASTAONDAGSTLNNAFNMGWGGQARWRARBOVAAQAAA 233
 321 DKFSRNADNOSFLVESVYAIN-----TIKAAVSPQNTINW---DKQLAGIYA- 365
 234 SDAGAV-----MSGIKVERNIVQTLILGGAVLADGKISAGAMTAGSTLVGRALAP 286
 366 --AGFKVTVLATIGQGQIQLIKFTVMINMLWGAHLVSGDLSIGQLAFNMLAGQIYAP 423
 287 IEGAVGQMKNTIGARGANDRLQTMARE-EKSAADHMPERPGRVLSAANAATLPPGAQOP 345
 424 VIRLAQIMQDQOVQVSVIRLGDVLSNPTESYHGKLAPEINGDI----- 468
 346 TMRQASFR-----IDAGAVALVLPSSAGKSLGLGIYGVMPCAAGVIRL 390
 469 TFRNIRFPRKDSVYIILDNILNLSIKQGEVIGIVGRSSGSKSTLTKLIDRFYIPENGOVLI 528
 391 DGYDIKQMDPEKLRHNGVYLPODIELEFSGTVAQNIARFTEES--OEVLTAATLGVHMI 449
 529 DGHDLAIDPMLRQGVVVLQDNVLNRSIINDNISLANPGMSVSKYIYAKKLAGAHDFI 588
 450 QSLPNGYTAIGEGGASISGQOROLARAVFRMPALLVDEPNVASIDQYGEVALMEAM 509
 589 SELREGYNTIIVGEGAGISGQORRIARALVNNPKILIFDEATSAIDYSEHVMNM 648
 510 KRLAKRTVIEFAPHKVNLAAQADYIMVINGVYSDPGE 548
 649 KKI--CKGRTVILIAHRLSTVKNADRIIVMERKGIYEGK 686

RESULT 9
 HLXB_ECOLI
 ID HLXB_ECOLI STANDARD: PRT: 707 AA.
 AC P08716; 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hemolysin secretion ATP-binding protein, plasmid.
 GN HLXB
 OS Escherichia coli.
 OG Plasmid Inc12 PHLY152.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hess J., Weis W., Vogel M., Goebel W.;
 RT "Nucleotide sequence of a plasmid-encoded hemolysin determinant and its comparison with a corresponding chromosomal hemolysin sequence.";
 RL FEBS Microbiol. Lett. 34:1-11(1986).
 RN [2]
 RP TOPOLOGY.
 RX MEDLINE=92204133; PubMed=1552901;
 RA Gentschev I., Goebel W.;
 RT "Topological and functional studies on HLXB of Escherichia coli.";
 RL Mol. Gen. Genet. 232:40-48(1992).
 RN [3]
 RP TOPOLOGY.
 RX MEDLINE=91132653; PubMed=1994034;
 RA Wang R.C., Seror S.J., Blight M., Pratt J.M., Broome-Smith J.K., Holland I.B.;
 RT "Analysis of the membrane organization of an Escherichia coli protein translocator, HLXB, a member of a large family of prokaryote and eukaryote surface transport proteins.";
 RL J. Mol. Biol. 217:441-454(1991).
 CC -1- FUNCTION: INVOLVED IN THE EXPORT OF HEMOLYSIN A.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. HLXB SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

Query Match 20.0%; Score 561.5; DB 1; Length 707;
 Best Local Similarity 29.7%; Pred. No. 2,1e-31;
 Matches 172; Conservative 101; Mismatches 233; Indels 73; Gaps 14;

| FT | DOMAIN | 409 | 411 | PERIPLASMTIC (PROBABLE). |
|----|----------|------------------------------------|--------|--------------------------|
| FT | TRANSMEM | 412 | 432 | PROBABLE. |
| FT | DOMAIN | 433 | 707 | CYTOPLASMTIC (PROBABLE). |
| FT | NP_BIND | 502 | 509 | ATP (BY SIMILARITY). |
| SO | SEQUENCE | 707 AA; 79672 MW; 4123ABE64A3CFEBA | CRC64; | |

Query Match 20.0%; Score 561.5; DB 1; Length 707;
 Best Local Similarity 29.7%; Pred. No. 2,1e-31;
 Matches 172; Conservative 101; Mismatches 233; Indels 73; Gaps 14;

| QY | 10 | TIFDQAVLVARPAVITAMVSEFINILALVSPMLQVYDRVLTSRNSTLIVTVICVF | 69 |
|----|-----|---|-----|
| DB | 141 | TWFPATIKYRRIRPIETLVSVFLQALFALITPLEFQVYMDKVLVHRGFSTLVNTVALSV | 200 |
| QY | 70 | LELVYGLLEALRTQVLVVRGKLFQVARDPIFKSVLDSTLS-----RKIGGQAFRDMQ | 124 |
| DB | 201 | VVVEIILSGLRITVIFSHSTRIDVELGAKLFRHLALPISEFSRRKGVAVARVRELDQ | 260 |
| QY | 125 | VREPMTG-GLIAFCDAWPVPVIVSWMLHP--FFGILAIACIIIFGLAV-----MN | 174 |
| DB | 261 | IRNFLTGALTSVLDLFLSLFFAVMWYSPKLTIVLIFSLPCYAAMSVFTSPILRRLD | 320 |
| QY | 175 | DNATKN-PIQMATMASTAAQDASTLNAEVMKAMGMLQARWRARRDQVAMQAAA | 233 |
| DB | 321 | DKFSRMDNOSFLVESVAIN-----TIKAMAVSPQMTNIM---DKQLAGYVA- | 365 |
| QY | 234 | SDAGAV-----MSGKVFERNIVQTLILGGAVALIDGKISAGAMTAGSILVGRALAP | 286 |
| DB | 366 | --AGKTVIVLATIGQGGIQLQKTYMIINLMGALHVLISGDLSTIGQLAFNMLAQIYAP | 423 |
| QY | 287 | IEGAVGQKNYIGARGADRLQTMRE-EKSAADHMLPPEPRGVLSAEASTLPPGAQP | 345 |
| DB | 424 | VIRLAQIMQDFQOQVIGISVTRGLDVLSNPTSEYHGKLTLPENIGDI----- | 468 |
| QY | 346 | TMRQASFR-----IDAGAVALVYPSAGKSSLLNGIYGVMPGACAVIRL | 390 |
| DB | 469 | TFRNIRFRKPDSPVILDNINLSIKQGEVIGIVRSGSGKSLTLILQRFYIPENGOYLI | 528 |
| QY | 391 | DGYIDKMDPEKLRHNGVLPQDIELFSGIVAAQNIARFTEPES--OEVIETATLAVHEMI | 449 |
| DB | 529 | DGHDLADLPNMLRQGVVVLQDNVNLNRSIIDNLSLNPMSGVKVIYAALACAGHPDI | 588 |
| QY | 450 | QSLPMGYTPTAIGEGGASISGSGORALARAFAVRPALIVDEPNASLDQGEVALMEM | 509 |
| DB | 589 | SELREGYNTIVIEGOGAGORHAIARALVNNPKILIFDEATSSALDYESEHIVRMN | 648 |
| QY | 510 | KRLKAARVVFATHKVNLLAQADYIMVINGVISDFGE | 548 |
| DB | 649 | HRI-CKGRVITIIAHRISTVKADRIYMEKGIVGQK | 686 |

RESULT 10
 RT3B_ACTPL STANDARD: PRT: 711 AA.
 ID RT3B_ACTPL 004473.
 AC 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE RTX-III toxin determinant B (Toxin RTX-III secretion ATP-binding protein) (APX-III) (Cytolysin IIIB) (CLY-IIIB).
 GN APXIII OR CLYIII OR RTX.
 OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Actinobacillus.
 OX NCBI_TaxID=715;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SEROTYPE 2;
 RX MEDLINE=9326392; PubMed=8494611;
 RA Chang Y.-F., Shi J.-Y., Ma D.-P., Shin S.-J., Lein D.H.;
 RT "Molecular analysis of the Actinobacillus pleuropneumoniae RTX toxin-III gene cluster."

Query Match 19.7%; Score 552.5; DB 1; Length 711;
 Best Local Similarity 28.9%; Pred. No. 8,7e-31;
 Matches 165; Conservative 117; Mismatches 238; Indels 51; Gaps 14;

| QY | 10 | TIFDQAVLVARPAVITAMVSEFINILALVSPMLQVYDRVLTSRNSTLIVTVICVF | 69 |
|----|-----|---|-----|
| DB | 145 | TWFPATIKYRRIRPIETLVSVFLQALFALITPLEFQVYMDKVLVHRGFSTLVNTVALSV | 204 |
| QY | 70 | LELVYGLLEALRTQVLVVRGKLFQVARDPIFKSVLDSTLS-----RKIGGQAFRDMQ | 124 |
| DB | 205 | VVVEIILSGLRITVIFSHSTRIDVELGAKLFRHLALPISEFENRNVGDTVARVRELDQ | 264 |

Query Match 19.7%; Score 552.5; DB 1; Length 711;
 Best Local Similarity 28.9%; Pred. No. 8,7e-31;
 Matches 165; Conservative 117; Mismatches 238; Indels 51; Gaps 14;

| RL | 121 | 351-362(1993). |
|----|--|---------------------------|
| RL | [12] | |
| RN | SEQUENCE FROM N.A. | |
| RP | STRAIN=405 / SEROTYPE 8; | |
| RC | MEDLINE=95012630; PubMed=7927703; | |
| RA | Jansen R., Briare J., van Geel A.B.M., Kamp E.M., Gielkens A.L.J., | |
| RA | Smits M.A.; | |
| RT | "Genetic map of the Actinobacillus pleuropneumoniae RTX-toxin (ApX) | |
| RT | operons: characterization of the ApXIII operons." | |
| RL | Infect. Immun. 62:441-4418(1994). | |
| RN | [3] | |
| RP | SEQUENCE OF 1-39 FROM N.A. | |
| RC | STRAIN=405 / SEROTYPE 8; | |
| RX | MEDLINE=93162836; PubMed=8432615; | |
| RA | Jansen R., Briare J., Kamp E.M., Gielkens A.L.J., Smits M.A.; | |
| RT | "Cloning and characterization of the Actinobacillus | |
| RT | pleuropneumoniae-RTX-toxin III (ApXIII) gene." | |
| RL | Infect. Immun. 61:947-954(1993). | |
| CC | -1- FUNCTION: INVOLVED IN THE TRANSPORT OF THE TOXIN RTX-III. | |
| CC | -1- SUBUNIT: HOMODIMER (BY SIMILARITY). | |
| CC | -1- SUBCELLULAR LOCATION: Integral membrane protein. | |
| CC | -1- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF SEROTYPE 2. | |
| CC | -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. HLVB SUBFAMILY. | |
| CC | ----- | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | |
| CC | use by non-profit institutions as long as its content is in no way | |
| CC | modified and this statement is not removed. Usage by and for commercial | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ | |
| CC | or send an email to license@sib-sib.ch). | |
| CC | ----- | |
| DR | EMBL; L12145; AAA21925.1; - | |
| DR | EMBL; X80055; CA56359.1; - | |
| DR | EMBL; X68815; CAA48712.1; ALT_INIT. | |
| DR | InterPro: IPR003593; AAA. | |
| DR | InterPro: IPR001140; ABC transporter_tmam. | |
| DR | InterPro: IPR003439; ABC transporter. | |
| DR | InterPro: IPR001687; ATP_GTP_A. | |
| DR | Pfam: PF00664; ABC_membrane; 1. | |
| DR | Pfam: SMO0005; ABC_tran; 1. | |
| DR | SMART: SMO0382; AAA; 1. | |
| DR | PROSITE: PS00211; ABC_TRANSPORTER; 1. | |
| KW | Cytolysin; Transport; ATP-binding; Transmembrane. | |
| FT | TRANSMEM 162 178 | POTENTIAL. |
| FT | TRANSMEM 197 213 | POTENTIAL. |
| FT | TRANSMEM 274 290 | POTENTIAL. |
| FT | TRANSMEM 299 315 | POTENTIAL. |
| FT | TRANSMEM 391 407 | POTENTIAL. |
| FT | TRANSMEM 506 513 | ATP (BY SIMILARITY). |
| FT | NP_BIND 61 61 | M -> A (IN SEROTYPE 8). |
| FT | VARIANT 91 91 | M -> T (IN SEROTYPE 8). |
| FT | VARIANT 125 126 | GM -> DV (IN SEROTYPE 8). |
| FT | VARIANT 213 213 | S -> N (IN SEROTYPE 8). |
| FT | VARIANT 416 416 | T -> A (IN SEROTYPE 8). |
| FT | VARIANT 469 469 | F -> K (IN SEROTYPE 8). |
| FT | VARIANT 606 606 | L -> Q (IN SEROTYPE 8). |
| FT | VARIANT 638 638 | S -> R (IN SEROTYPE 8). |
| SO | SEQUENCE 711 AA; 80405 MW; 5B6E46B89D12D92 CRC64; | |

| | | | |
|----|-----|--|-----|
| Qy | 125 | VREFFETG-GLLAFCDAPWTPEVIVSNMLHPFECIALIA--CIIFGLAVMNDNATKNP | 181 |
| Db | 265 | IRNFLTGALTSVDLDFESFTFEFVMMYTSPKLTIIVLLISLPTIAMSTI-----SP | 317 |
| Qy | 182 | IQMTMASIAAQN-DAGSTLRNA---EYKAMGMGGLQARMRARDEQVAAQASDA | 236 |
| Db | 318 | ILRRRLDEKFRANDNDOSFLVESASADITKLALVATPQMNINW---DKOLASVSA-DF | 372 |
| Qy | 217 | GGAVMS-----GIVFENIVOTLLGGGAVIATIDKISAGAMTAGSTIVRATALPIEGAV | 291 |
| Db | 373 | RYIVLATITGGQGVLLDRITVMTIIMLGMGAHVLSGDISTQGLTFNMLSGQVIAVPVRLA | 432 |
| Qy | 292 | GOMKNITGARGAMDLOTMLRE-EKSDADHMPLEPRGVLS-----AEAASILPG | 341 |
| Db | 433 | QLMQDFQGVGISTIRLDDVLSNPTEYNQKLSLPEIGDLAFKHIRRYRKPDAPIILD-- | 490 |
| Qy | 342 | AQQPTMQASRIDAGAAVAALVYPSAAGKSSLLRGIVGWPCCAGVIRLDGTDIKOMPE | 401 |
| Db | 491 | -----DVNLSTVQGEVYIGIVSGSGSKETPLKLLDRFTIPENGQVLDIGHDLALADPN | 543 |
| Qy | 402 | KLGRHVGVLPODIELFSGTVAQNTARFTEPESQE-VIEATYTLAGVHMIOSLPRGYDTAI | 460 |
| Db | 544 | WLRQIGVLDQDNVLLNRSIRDNALITLDPKSMSPRYIYAKKLAGANDFISELDEGVIYIV | 603 |
| Qy | 461 | GEGGASLISGGORRLALARAFAFMPPALLVLDEPNASILDQVEVALMEAMRLKAARTVI | 520 |
| Db | 604 | GELGAGLSGGGQRRIATARALVNNPRILIFDEPATSALDYESEHIMQNKI-CHGRTVI | 662 |
| Qy | 521 | FATHKVNLDAQDYIMVINGVTSDEEQROR | 551 |
| Db | 663 | IIARLSTVIGKNADRIIYMGEGHIVLEQGNKK | 693 |

| RESULT | 11 |
|------------|--|
| CYAB_BORPE | |
| ID | CYAB_BORPE |
| AC | P18770; |
| DT | 01-NOV-1990 (Rel. 16, Created) |
| DT | 01-NOV-1990 (Rel. 16, Last sequence update) |
| DT | 16-OCT-2001 (Rel. 40, Last annotation update) |
| DE | Cyclolysin secretion ATP-binding protein cyab. |
| GN | CYAB. |
| OS | Bordetella pertussis. |
| OC | Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae; |
| OC | Bordetella. |
| OX | NCBI_TaxID=520; |
| RN | [1] |
| RP | SEQUENCE FROM N.A. |
| RC | STRAIN=18323; |
| RX | MEDLINE=69091151; PubMed=2905265; |
| RA | Glaser P., Sakamoto H., Bellalou J., Uilmann A., Danchin A.; |
| RT | "Secretion of cyclolysin, the calmodulin-sensitive adenylate cyclase- |
| RL | haemolysin bifunctional protein of Bordetella pertussis."; |
| CC | EMBO J. 7:3997-4004(1988). |
| CC | -1- FUNCTION: INVOLVED IN THE EXPORT OF CALMODULIN-SENSITIVE |
| CC | ADENYLATE CYCLASE-HEMOLYSIN (CYCLOLYSIN). |
| CC | -1- SUBCELLULAR LOCATION: Integral membrane protein. |
| CC | -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. HLYB SUBFAMILY. |
| CC | ----- |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - |
| CC | the European Bioinformatics Institute. There are no restrictions on its |
| CC | use by non-profit institutions as long as its content is in no way that of |
| CC | modified and this statement is not removed. Usage by and for commercial |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ |
| CC | or send an email to license@sib-sib.ch). |
| CC | ----- |
| DR | EMBL; X14199; CAA32412.1; - |
| DR | PIR; S02386; BVBRCB. |
| DR | InterPro; IPR003593; AAA. |
| DR | InterPro; IPR001140; ABC_transporter_tmem. |
| DR | InterPro; IPR003439; ABC_transportr. |
| DR | InterPro; IPR001687; ATP_GTP_A. |

| | |
|-----|---|
| DR | Pfam: PF000664; ABC_membrane; 1. |
| DR | Pfam: PF00005; ABC_tran; 1. |
| DR | SMART; SM00382; AAA; 1. |
| DR | PROSITE; PS00211; ABC_TRANSPORTER; 1. |
| KW | Hemolysis; Transp; ATP-binding; Transmembrane. |
| FT | TRANSMEM 105 122 POTENTIAL. |
| FT | TRANSMEM 226 242 POTENTIAL. |
| FT | TRANSMEM 261 282 POTENTIAL. |
| FT | TRANSMEM 295 315 POTENTIAL. |
| FT | TRANSMEM 368 392 POTENTIAL. |
| FT | TRANSMEM 399 419 POTENTIAL. |
| FT | TRANSMEM 464 488 POTENTIAL. |
| FT | NP_BIND 505 512 ATP (BY SIMILARITY). |
| SEQ | SEQUENCE 712 AA; 77969 MW; BA8DDC1430960675 CRC64 |

| | | | | |
|-----------------------|-------------------|--------------------|------------|-------------|
| Query Match | 19.0%; | Score 533; | DB 1; | Length 712; |
| Best Local Similarity | 27.0%; | Pred. NO. 1.9e-29; | | |
| Matches 166; | Conservative 105; | Mismatches 247; | Indels 96; | Gaps 15; |

| | | | | |
|----|-----|--|--------------------------|-----|
| Qy | 13 | DOAVAVAPAVITAAV----- | -SSEF----- | 33 |
| | | : | | |
| Db | 98 | DOAVLIGRQGAAPARLGAEPFALMAGELLACACASPQTQALANFDPSEWIFALVKRHLL | | 157 |
| Qy | 34 | -----NIALVSLPMLQYDPVVLTSRNVSTLYTVCFLFLYTGLELRT | | 82 |
| | | : | : | |
| Db | 158 | IGEVLLLSLVQTFISLPLPEFQYVMDVAVLNNAEETLNTVTVFLLALFEALLTGIRT | | 217 |
| Qy | 83 | QVLVVGGLKFDGVARDPIEKSVDLSTLS---- | KKIGGQAFRMDQYREEMGGLI-AF | 136 |
| | | : | : | |
| Db | 218 | YLFATHTSKLDELVGARLYVAHLRLPLAVYQFARRGGGSVARELEHIAKFLVGNAAVTVL | | 277 |

| | | | | | | |
|----|----|-----|--|---|-------------|-----|
| 27 | DB | 278 | LDVFSVFVFIAMF---- | FYSVKLTLYVLLALPCYFLISTLV----- | TPVLRRLRHVK | 326 |
| 28 | QY | 188 | ASIAQAAND--GSTLRNAEVRKAMGIMMGLOARRRARDEOVAMQAASADGAVMSGIK | | 245 | |
| 29 | DB | 327 | FNRCGENAFLEVEYSGLDIYKSL----- | AVEPQOMRMDRLAGIYVA----- | GLSTAVNA | 378 |
| 30 | QY | 246 | VERNIYQTLIG----- | GGAYLAIADGKISAGAMIAGSILVGRALAPIEGAVQWKNYI | | 298 |
| 31 | DB | 379 | MLANGVTLLIRLLRWESCAGVHNGRARRMYGETELVAFNMLSGHVTPVIRLAQIMNDFQ | | 438 | |
| 32 | QY | 299 | GARGAMDLOQMRLR--EESASDDHNPLEPRPREVLAEAAASILPRBAQOPTRKQASFRIDAG | | 357 | |
| 33 | DB | 439 | QTVSMQGLGILNLCRTFVADKAQLPALRKSIELDRVSFRYRPDAADALRNYSRLIAPG | | 498 | |
| 34 | QY | 358 | AAVALVGRSAKSSLLNGIYVWPACAGVIRLGGYDILKQMDPREKLRHNGYIYDODILEF | | 417 | |
| 35 | DB | 499 | EYVGVGVSNGSGSKTLTRILORMYVADRGVYLDIGHDILGYDSASLKRQOLAVYLQESTLF | | 558 | |
| 36 | QY | 418 | SGTYAONARTFTEES--OEVTEAATLAGVHEMIOSLPMGYDTAIEGASISGGOROLA | | 476 | |
| 37 | DB | 559 | NRSVADNIALTRPGASMHVEVAAARLGAHBEFIQDLPEGVDTMGENGVSGSQRRIG | | 618 | |
| 38 | QY | 477 | LARAVFRPALLVDEPNASIDQYGEVALMEAMKRLKAARTVIFATHKVNLQAADYIM | | 536 | |
| 39 | DB | 619 | IARALIHPRPVLILDEAFASALDYESEHIORNRMDI--CDGRTVILIIHRLSAVACADRIY | | 677 | |
| 40 | QY | 537 | VINGVVISDPEGRD | 550 | | |
| 41 | DB | 678 | WMEGGEVAECGSH | 691 | | |

| RESULT | 12 |
|---|-------------------------|
| ABC6_HUMAN | |
| ABC6_HUMAN | STANDARD; |
| Q9NP58: 075542; | PRT: 842 AA. |
| 16-OCT-2001 (Rel. 40, Created) | |
| 16-OCT-2001 (Rel. 40, Last sequence update) | |
| 16-OCT-2001 (Rel. 40, Last annotation update) | |
| Apr-binding cassette, sub-family B, member 6, | mitochondrial precursor |

| | | |
|----|--|-----|
| DR | PDB: 1J50; 12-SEP-01. | |
| DR | Ecogène; EGI0613; msbA. | |
| DR | InterPro: IPR003593; AAA. | |
| DR | InterPro: IPR001140; ABC_transporter_tnem. | |
| DR | InterPro: IPR003439; ABC_transportr. | |
| DR | InterPro: IPR001687; ATP_Grp_A. | |
| DR | Pfam: PF00664; ABC_membrane; 1. | |
| DR | Pfam: PF00005; ABC_tran; 1. | |
| DR | SMART: SM00382; AAA; 1. | |
| DR | PROSITE: PS00211; ABC_TRANSPORTER; 1. | |
| KW | ATP-binding; Transport; Inner membrane; Transmembrane; 3D-structure; | |
| RV | Complete proteome. | |
| FT | DOMAIN | 1. |
| FT | DOMAIN | 21 |
| FT | DOMAIN | 22 |
| FT | DOMAIN | 53 |
| FT | DOMAIN | 64 |
| FT | DOMAIN | 65 |
| FT | DOMAIN | 97 |
| FT | DOMAIN | 139 |
| FT | DOMAIN | 140 |
| FT | DOMAIN | 164 |
| FT | DOMAIN | 165 |
| FT | DOMAIN | 167 |
| FT | DOMAIN | 168 |
| FT | DOMAIN | 192 |
| FT | DOMAIN | 193 |
| FT | DOMAIN | 252 |
| FT | DOMAIN | 253 |
| FT | DOMAIN | 272 |
| FT | DOMAIN | 273 |
| FT | DOMAIN | 280 |
| FT | DOMAIN | 281 |
| FT | DOMAIN | 301 |
| FT | DOMAIN | 302 |
| FT | DOMAIN | 376 |
| FT | DOMAIN | 383 |
| FT | DOMAIN | 384 |
| FT | DOMAIN | 385 |
| FT | DOMAIN | 386 |
| FT | DOMAIN | 387 |
| FT | DOMAIN | 388 |
| FT | DOMAIN | 389 |
| FT | DOMAIN | 390 |
| FT | DOMAIN | 391 |
| FT | DOMAIN | 392 |
| FT | DOMAIN | 393 |
| FT | DOMAIN | 394 |
| FT | DOMAIN | 395 |
| FT | DOMAIN | 396 |
| FT | DOMAIN | 397 |
| FT | DOMAIN | 398 |
| FT | DOMAIN | 399 |
| FT | DOMAIN | 400 |
| FT | DOMAIN | 401 |
| FT | DOMAIN | 402 |
| FT | DOMAIN | 403 |
| FT | DOMAIN | 404 |
| FT | DOMAIN | 405 |
| FT | DOMAIN | 406 |
| FT | DOMAIN | 407 |
| FT | DOMAIN | 408 |
| FT | DOMAIN | 409 |
| FT | DOMAIN | 410 |
| FT | DOMAIN | 411 |
| FT | DOMAIN | 412 |
| FT | DOMAIN | 413 |
| FT | DOMAIN | 414 |
| FT | DOMAIN | 415 |
| FT | DOMAIN | 416 |
| FT | DOMAIN | 417 |
| FT | DOMAIN | 418 |
| FT | DOMAIN | 419 |
| FT | DOMAIN | 420 |
| FT | DOMAIN | 421 |
| FT | DOMAIN | 422 |
| FT | DOMAIN | 423 |
| FT | DOMAIN | 424 |
| FT | DOMAIN | 425 |
| FT | DOMAIN | 426 |
| FT | DOMAIN | 427 |
| FT | DOMAIN | 428 |
| FT | DOMAIN | 429 |
| FT | DOMAIN | 430 |
| FT | DOMAIN | 431 |
| FT | DOMAIN | 432 |
| FT | DOMAIN | 433 |
| FT | DOMAIN | 434 |
| FT | DOMAIN | 435 |
| FT | DOMAIN | 436 |
| FT | DOMAIN | 437 |
| FT | DOMAIN | 438 |
| FT | DOMAIN | 439 |
| FT | DOMAIN | 440 |
| FT | DOMAIN | 441 |
| FT | DOMAIN | 442 |
| FT | DOMAIN | 443 |
| FT | DOMAIN | 444 |
| FT | DOMAIN | 445 |
| FT | DOMAIN | 446 |
| FT | DOMAIN | 447 |
| FT | DOMAIN | 448 |
| FT | DOMAIN | 449 |
| FT | DOMAIN | 450 |
| FT | DOMAIN | 451 |
| FT | DOMAIN | 452 |
| FT | DOMAIN | 453 |
| FT | DOMAIN | 454 |
| FT | DOMAIN | 455 |
| FT | DOMAIN | 456 |
| FT | DOMAIN | 457 |
| FT | DOMAIN | 458 |
| FT | DOMAIN | 459 |
| FT | DOMAIN | 460 |
| FT | DOMAIN | 461 |
| FT | DOMAIN | 462 |
| FT | DOMAIN | 463 |
| FT | DOMAIN | 464 |
| FT | DOMAIN | 465 |
| FT | DOMAIN | 466 |
| FT | DOMAIN | 467 |
| FT | DOMAIN | 468 |
| FT | DOMAIN | 469 |
| FT | DOMAIN | 470 |
| FT | DOMAIN | 471 |
| FT | DOMAIN | 472 |
| FT | DOMAIN | 473 |
| FT | DOMAIN | 474 |
| FT | DOMAIN | 475 |
| FT | DOMAIN | 476 |
| FT | DOMAIN | 477 |
| FT | DOMAIN | 478 |
| FT | DOMAIN | 479 |
| FT | DOMAIN | 480 |
| FT | DOMAIN | 481 |
| FT | DOMAIN | 482 |
| FT | DOMAIN | 483 |
| FT | DOMAIN | 484 |
| FT | DOMAIN | 485 |
| FT | DOMAIN | 486 |
| FT | DOMAIN | 487 |
| FT | DOMAIN | 488 |
| FT | DOMAIN | 489 |
| FT | DOMAIN | 490 |
| FT | DOMAIN | 491 |
| FT | DOMAIN | 492 |
| FT | DOMAIN | 493 |
| FT | DOMAIN | 494 |
| FT | DOMAIN | 495 |
| FT | DOMAIN | 496 |
| FT | DOMAIN | 497 |
| FT | DOMAIN | 498 |
| FT | DOMAIN | 499 |
| FT | DOMAIN | 500 |
| FT | DOMAIN | 501 |
| FT | DOMAIN | 502 |
| FT | DOMAIN | 503 |
| FT | DOMAIN | 504 |
| FT | DOMAIN | 505 |
| FT | DOMAIN</ | |

Search completed: July 19, 2002, 10:49:04
Job time: 404 sec

| | | | | | | | |
|-----------------------|-------|--------------|--------|--------|-----|--------|------|
| Query Match | 15.7% | Score | 440; | DB | 1; | length | 582; |
| Best Local Similarity | 26.7% | Pred. No. | 4e-23; | | | | |
| Matches | 152; | Conservative | 248; | Indels | 58; | Gaps | 17; |

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2002, 10:41:45 ; Search time 53.06 Seconds
(without alignments)
1809.502 Million cell updates/sec

Title: US-09-913-414-4

Perfect score: 2807

Sequence: 1 MFKRSQAKPTFDQAVLVAR.....MVINQGVISDFGERDQWPS 555

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_19:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_protent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rviro:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1 | 2765 | 98.5 | 578 | 16 | 085350 |
| 2 | 2759 | 98.3 | 578 | 2 | 09RMNO |
| 3 | 1206.5 | 43.0 | 583 | 2 | 09Z3C5 |
| 4 | 1195 | 42.6 | 580 | 2 | 09RHR2 |
| 5 | 1193.5 | 42.5 | 583 | 2 | 09ZNT9 |
| 6 | 1184 | 42.2 | 578 | 2 | 09ZG94 |
| 7 | 1182.5 | 42.1 | 578 | 2 | 054456 |
| 8 | 1177.5 | 41.9 | 579 | 2 | 067993 |
| 9 | 1176.5 | 41.9 | 588 | 2 | 09XB64 |
| 10 | 1172 | 41.8 | 583 | 2 | 087808 |
| 11 | 1164.5 | 41.5 | 599 | 2 | 09KGS6 |
| 12 | 1156 | 41.2 | 600 | 16 | 09HYJ8 |
| 13 | 1138 | 40.5 | 582 | 2 | 054416 |
| 14 | 1088.5 | 38.8 | 617 | 16 | 092MX3 |
| 15 | 1077 | 38.4 | 582 | 16 | 033678 |
| 16 | 1068 | 38.0 | 581 | 16 | 098L68 |

| | | | | | | |
|----|--------|------|------|----|---------|---------------------|
| 17 | 1060.5 | 37.8 | 570 | 16 | 0988E6 | 0988E6 rhizobium l |
| 18 | 1052.5 | 37.5 | 571 | 2 | 09X6N6 | 096n6 rhizobium l |
| 19 | 1044.5 | 37.2 | 564 | 16 | 0677184 | 0677184 aquilex aeo |
| 20 | 1039 | 37.0 | 576 | 2 | 053368 | 053368 serratia ma |
| 21 | 1039 | 37.0 | 586 | 2 | 054451 | 054451 serratia ma |
| 22 | 1038 | 37.0 | 570 | 2 | 057040 | 057040 rhizobium l |
| 23 | 1036.5 | 36.9 | 579 | 2 | 085375 | 085375 proteus mlr |
| 24 | 1033 | 36.8 | 570 | 2 | 005693 | 005693 rhizobium l |
| 25 | 1012 | 36.1 | 583 | 16 | 09ZDL3 | 09ZDL3 rickettsia |
| 26 | 1011 | 36.0 | 584 | 16 | 09Z1J1 | 09Z1J1 rickettsia |
| 27 | 958 | 34.1 | 589 | 2 | P66439 | P66439 rhizobium m |
| 28 | 958 | 34.1 | 589 | 16 | 092V33 | 092V33 rhizobium m |
| 29 | 957 | 34.1 | 589 | 2 | 09X7L7 | 09X7L7 rhizobium m |
| 30 | 913 | 32.5 | 712 | 16 | 092LS9 | 092LS9 rhizobium m |
| 31 | 892 | 31.8 | 581 | 16 | 098139 | 098139 rhizobium m |
| 32 | 596.5 | 21.3 | 706 | 2 | 047462 | 047462 escherichia |
| 33 | 592.5 | 21.1 | 706 | 2 | 046717 | 046717 escherichia |
| 34 | 590.5 | 21.0 | 705 | 2 | 09LC57 | 09LC57 escherichia |
| 35 | 582.5 | 20.8 | 708 | 2 | 093FG6 | 093FG6 pasteurella |
| 36 | 581.5 | 20.7 | 708 | 2 | 093FH0 | 093FH0 pasteurella |
| 37 | 580.5 | 20.7 | 708 | 2 | 093FH3 | 093FH3 pasteurella |
| 38 | 580.5 | 20.7 | 708 | 2 | 0934E0 | 0934E0 pasteurella |
| 39 | 580.5 | 20.7 | 708 | 2 | 0934A3 | 0934A3 pasteurella |
| 40 | 577.5 | 20.6 | 708 | 2 | 093FG8 | 093FG8 pasteurella |
| 41 | 577.5 | 20.6 | 708 | 2 | 093FG4 | 093FG4 mannheimia |
| 42 | 577.5 | 20.6 | 708 | 2 | 093FG2 | 093FG2 mannheimia |
| 43 | 577.5 | 20.6 | 708 | 2 | 0933I3 | 0933I3 mannheimia |
| 44 | 577.5 | 20.6 | 1011 | 16 | P74176 | P74176 synechocyst |
| 45 | 576.5 | 20.5 | 708 | 2 | 093FH6 | 093FH6 pasteurella |

ALIGNMENTS

RESULT 1

ID 085350 PRELIMINARY: PRT: 578 AA.

AC 085350:

DT 01-NOV-1998 (TREMBLrel. 08, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE ABC TRANSPORTER (RSAA SECRETION SYSTEM, ATP-BINDING PROTEIN RSAD).

GN RSAD OR CCI1008.

OS Caulobacter crescentus.

OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;

OC Caulobacter.

OX NCBI_TaxID=69394;

RN [1]

RP SEQUENCE OF 1-98 FROM N.A.

RC STRAIN-ATCC 19089 / CB15, AND NA1000;

RX MEDLINE=85054630; PubMed=6209263;

RA Smit J.K., Agabian N.;

RT "Cloning of the major protein of the Caulobacter crescentus periodic surface layer: detection and characterization of the cloned peptide by protein expression assays.";

RT J. Bacteriol. 160:1137-1145(1984).

RL [2]

RN SEQUENCE FROM N.A.

RP STRAIN-ATCC 19089 / CB15, AND NA1000;

RC MEDLINE=98292737; PubMed=9620954;

RA Avram P., Smit J.;

RT "The Caulobacter crescentus paracrystalline S-layer protein is secreted by an ABC transporter (Type I) secretion apparatus.";

RL J. Bacteriol. 180:3062-3069(1998).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN-ATCC 19089 / CB15, AND NA1000;

RA Avram P., Smit J.K.;

RT Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.

RL [4]

RP SEQUENCE FROM N.A.

RC STRAIN-ATCC 19089 / CB15, AND NA1000;

RA Avram P.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC SRRAIN-ATCC 19089 / CB15;
RX MEDLINE-21173698; PubMed-11259647;
RA Nierman W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Elesen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utechtack T., Tran K., Wolf A., Yamathayan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.,
RT "Complete genome sequence of *Caulobacter crescentus*.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
(ABC TRANSPORTERS).
CC EMBL: AF062345; AAC38666.2; -;
DR EMBL: AE005779; AAK22992.1; -;
DR HSSP: P13569; INBD.
DR TIGR: CC1008; -;
DR InterPro: IPR001140; ABC_transporter_tmam.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam: PF00664; ABC_membrane; 1.
DR Pfam: PF00005; ABC_tran; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding; Complete proteome; Transport.
SQ SEQUENCE 578 AA; 62019 MW; 1E53C2D4A97675C5 CRC64;

Query Match 98.5%; Score 2765; DB 16; Length 578;
Best Local Similarity 99.8%; Pred. No. 4.9e-160;
Matches 549; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MFKRGARPTFDQAVLVARPAVITAMVFSFPIINITALVSPLYMLOYVDRVITSRNVSTL 60
DB 1 MFKRGARPTFDQAVLVARPAVITAMVFSFPIINITALVSPLYMLOYVDRVITSRNVSTL 60
OY 61 IYLVYICVFLVYGLLELRTQVLRGGLKPDGVARDPFEKSVSDSTLSRKGIGGOAFR 120
DB 61 IYLVYICVFLVYGLLELRTQVLRGGLKPDGVARDPFEKSVSDSTLSRKGIGGOAFR 120
OY 121 DMDQVREFMTGGLIAFCDAWPTPVFIYSWMLHPFEGILAIITACIIIFGLAVMNDATKN 180
DB 121 DMDQVREFMTGGLIAFCDAWPTPVFIYSWMLHPFEGILAIITACIIIFGLAVMNDATKN 180
OY 121 DMDQVREFMTGGLIAFCDAWPTPVFIYSWMLHPFEGILAIITACIIIFGLAVMNDATKN 180
DB 121 DMDQVREFMTGGLIAFCDAWPTPVFIYSWMLHPFEGILAIITACIIIFGLAVMNDATKN 180
OY 181 PIOMATMASIAQNDAGSTLRNAEYKAMGMGLOARWRARDEQVANOAAASDAGAV 240
DB 181 PIOMATMASIAQNDAGSTLRNAEYKAMGMGLOARWRARDEQVANOAAASDAGAV 240
OY 241 MSGIKVFRNIYOTLLIGGAYLAIDCKISAGAMINGSLVGRALPIEGAVGOMKNYIGA 300
DB 241 MSGIKVFRNIYOTLLIGGAYLAIDCKISAGAMINGSLVGRALPIEGAVGOMKNYIGA 300
OY 241 MSGIKVFRNIYOTLLIGGAYLAIDCKISAGAMINGSLVGRALPIEGAVGOMKNYIGA 300
DB 241 MSGIKVFRNIYOTLLIGGAYLAIDCKISAGAMINGSLVGRALPIEGAVGOMKNYIGA 300
OY 301 RGAMDRLOTMLREKESADDMPLPEPRGVLSAEASILPPGAQOQPTMRQASFRIDAGAAV 360
DB 301 RGAMDRLOTMLREKESADDMPLPEPRGVLSAEASILPPGAQOQPTMRQASFRIDAGAAV 360
OY 361 ALVGPAAKSSLLRGIVGWPCAGVIRLDGDIKOMPBEKIGRHVGLPODIELEFSGT 420
DB 361 ALVGPAAKSSLLRGIVGWPCAGVIRLDGDIKOMPBEKIGRHVGLPODIELEFSGT 420
OY 421 VAONIRARFEFESQEVIEATLAGVHEMIQSLPMGYDTAIGEGASLSGGORRLALARA 480
DB 421 VAONIRARFEFESQEVIEATLAGVHEMIQSLPMGYDTAIGEGASLSGGORRLALARA 480
OY 481 VFRMPALVLDEPNASLDQVGEVALMEAMKRLKAARKTVIFATHKNLLAQAADYIMVINO 540
DB 481 VFRMPALVLDEPNASLDQVGEVALMEAMKRLKAARKTVIFATHKNLLAQAADYIMVINO 540
OY 541 GVISDFGERD 550
DB 541 GVISDFGERD 550

DB 541 GVISDFGERD 550

RESULT 2
O99RNO PRELIMINARY; PRT; 578 AA.
ID O99RNO;
AC O99RNO;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ABC_TRANSPORTER.
GN RSAD.
OS *Caulobacter crescentus*.
OC Bacteria; Proteobacteria; alpha subdivision; *Caulobacter* group;
OC *Caulobacter*.
OX NCBI_TaxID=69394;
RN [1]
RP SEQUENCE FROM N.A.
RC SRRAIN-J54000;
RA Bingle W.H., Avram P.A., Nomellini J.F., Smit J.K.;
RT "The secretion signal of the *Caulobacter crescentus* S-layer protein is
RT located within the C-terminal 82 amino acids of the molecule.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE ABC_TRANSPORTER FAMILY.
DR EMBL: AF193064; AAF07962.1; -;
DR HSSP: P13569; INBD.
DR InterPro: IPR001140; ABC_transporter_tmam.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam: PF00664; ABC_membrane; 1.
DR Pfam: PF00005; ABC_tran; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding; Transport.
SQ SEQUENCE 578 AA; 62006 MW; 8661C449E0A47ADB CRC64;

Query Match 98.3%; Score 2759; DB 2; Length 578;
Best Local Similarity 99.6%; Pred. No. 1.1e-159;
Matches 548; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 MFKRGARPTFDQAVLVARPAVITAMVFSFPIINITALVSPLYMLOYVDRVITSRNVSTL 60
DB 1 MFKRGARPTFDQAVLVARPAVITAMVFSFPIINITALVSPLYMLOYVDRVITSRNVSTL 60
OY 61 IYLVYICVFLVYGLLELRTQVLRGGLKPDGVARDPFEKSVSDSTLSRKGIGGOAFR 120
DB 61 IYLVYICVFLVYGLLELRTQVLRGGLKPDGVARDPFEKSVSDSTLSRKGIGGOAFR 120
OY 121 DMDQVREFMTGGLIAFCDAWPTPVFIYSWMLHPFEGILAIITACIIIFGLAVMNDATKN 180
DB 121 DMDQVREFMTGGLIAFCDAWPTPVFIYSWMLHPFEGILAIITACIIIFGLAVMNDATKN 180
OY 121 DMDQVREFMTGGLIAFCDAWPTPVFIYSWMLHPFEGILAIITACIIIFGLAVMNDATKN 180
DB 121 DMDQVREFMTGGLIAFCDAWPTPVFIYSWMLHPFEGILAIITACIIIFGLAVMNDATKN 180
OY 181 PIOMATMASIAQNDAGSTLRNAEYKAMGMGLOARWRARDEQVANOAAASDAGAV 240
DB 181 PIOMATMASIAQNDAGSTLRNAEYKAMGMGLOARWRARDEQVANOAAASDAGAV 240
OY 241 MSGIKVFRNIYOTLLIGGAYLAIDCKISAGAMINGSLVGRALPIEGAVGOMKNYIGA 300
DB 241 MSGIKVFRNIYOTLLIGGAYLAIDCKISAGAMINGSLVGRALPIEGAVGOMKNYIGA 300
OY 241 MSGIKVFRNIYOTLLIGGAYLAIDCKISAGAMINGSLVGRALPIEGAVGOMKNYIGA 300
DB 241 MSGIKVFRNIYOTLLIGGAYLAIDCKISAGAMINGSLVGRALPIEGAVGOMKNYIGA 300
OY 301 RGAMDRLOTMLREKESADDMPLPEPRGVLSAEASILPPGAQOQPTMRQASFRIDAGAAV 360
DB 301 RGAMDRLOTMLREKESADDMPLPEPRGVLSAEASILPPGAQOQPTMRQASFRIDAGAAV 360
OY 361 ALVGPAAKSSLLRGIVGWPCAGVIRLDGDIKOMPBEKIGRHVGLPODIELEFSGT 420
DB 361 ALVGPAAKSSLLRGIVGWPCAGVIRLDGDIKOMPBEKIGRHVGLPODIELEFSGT 420
OY 421 VAONIRARFEFESQEVIEATLAGVHEMIQSLPMGYDTAIGEGASLSGGORRLALARA 480
DB 421 VAONIRARFEFESQEVIEATLAGVHEMIQSLPMGYDTAIGEGASLSGGORRLALARA 480
OY 481 VFRMPALVLDEPNASLDQVGEVALMEAMKRLKAARKTVIFATHKNLLAQAADYIMVINO 540
DB 481 VFRMPALVLDEPNASLDQVGEVALMEAMKRLKAARKTVIFATHKNLLAQAADYIMVINO 540

Db 481 VFRPALVYDEPNASLDQGEVALMEAMKRLKAARVTFATHKNVLLAQADYIMVING 540
OY 541 GVISDFGERD 550
Db 541 GVISDFGERD 550

RESULT 3

OY923C5 PRELIMINARY: PRT; 583 AA.
AC 0923C5;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE SAPP.
GN SAPP.
OS Campylobacter fetus.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=196;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=84-107, AND 23D;
RX MEDLINE=99069317; PubMed=9851986;
RA Thompson S.A., Shedd O.L., Ray K.C., Belns M.H., Jorgensen J.P.,
Blaser M.J.;
RT "Campylobacter fetus surface layer proteins are transported by a type
I secretion system";
RT J. Bacteriol. 180:6450-6458(1998).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL: AF071883; AAC97197.1; -;
DR EMBL: AF027405; AAC97164.1; -;
DR HSPF: P13569; INED.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR001140; ABC_transporter_tmam.
DR InterPro: IPR003439; ABC_transporter.
DR Pfam: PF00664; ABC_membrane.1.
DR Pfam: PF00005; ABC_tran.1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
DR ATP-binding; Transport.
SQ SEQUENCE 583 AA; 63958 MW; 891C8046136B7FAB CRC64;

Query Match 43.0%; Score 1206.5; DB 2; Length 583;
Best Local Similarity 45.5%; Pred. No. 1.9e-65;
Matches 245; Conservative 110; Mismatches 181; Indels 3; Gaps 3;

OY 14 QAVLVARPAVITAMVSPFINITLALVSPLYMLOYDRVLTSRNVSTLIVTCVFLFLIV 73
Db 13 ETKKSKNCIVTAITSFVFNLMILTPPLMLOLYRVVTSKESSTLFLTLVLFLEFTR 72
OY 74 YGLLEALRTQVLRGGLKFDGVARDPIFKSVLDSTLSRKSG-IGQAFRMDQVREFM-TG 131
Db 73 MAIFELRSRIILVFENQDIDLNTDRIYDAIFKLARSYGRVTSQMGDLNATKQYMNSTN 132
OY 132 GLIACFADFWTPEVTVVSMNLHPFEGILAIACIIIFGLAVNMNDATKNPIOMATYASIA 191
Db 133 GJFAFLDAFWLPLYIAILFEVHFWMGFIAAAILFGLVALLNEKATKDKLKSNDPTYKN 192
OY 192 AQNDASTLRNMEVNMKMGWGLOARWRARREQVAMQAAASDAGAVMSGIKYFRNIV 251
Db 193 EMRLIDMNLNSSEVINAMGNNNNLKIKMKHHTFNLNSHSDASIKAGVYTNISKATRVTS 252
OY 252 QTLILGGAYLADIKTISAGAMTAGSILVGRALAPIEGAVGQKNYIGARGMDRLQTM 311
Db 253 QSMGLGLGAYLVAKMEVSGMMIAGSILGRALAPIDILASKSKYKNRKESTERDKFL 312
OY 312 REKSKADHMLPLPEPVGIVSAEASILPEGAQOPTMRQASFRIDAGAAVALVPSAAGS 371
Db 313 HDEPVEKDLISLPDPRGDIACEAISLIPPSAKQPSLIGVSFALDADDMCAIIGPSAAGS 372

OY 372 SLRIGVWPCAGVYRLDGYDIKQWDPKLGHNVTLPDILEFSGVTAQNIARFTEF 431
Db 373 SLARALIGIMPAHGVYRVDCADINQYSDALGFPVYTLQDVELFGSTIAENIARGEL 432
OY 432 ESOQVIEATLACVHEMISQSLPMGYDTAIGEGASISGGQRQRLATARAVERMPALLVLD 491
Db 433 DSAVAVEAKSANNVHDMILRLPDGYDTKIGLGMSSISGGQRQRLALARKKPKIIVLD 492
OY 492 EPNASLDQGEVALMEAMKRLKAARVTFATHKNVLLAQADYIMVINGVISDFGERD 550
Db 493 EPNASLDEEGERALYALLAMK-GKATITLITRKNLVLAQVADKIAVLAMQOLVYFGERD 550

RESULT 4

OY9RHT2 PRELIMINARY: PRT; 580 AA.
AC 09RHT2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ABC PROTEIN.
GN HASD.
OS Pseudomonas fluorescens.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=294;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NO.33;
RX MEDLINE=20069636; PubMed=10601212;
RA Idel A., Kawai E., Akatsuka H., Omori K.;
RT "Cloning and characterization of the Pseudomonas fluorescens ATP-
binding cassette exporter, HasDEF, for the heme acquisition protein
HasA";
RT J. Bacteriol. 181:7545-7551(1999).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL: AB023289; BAA8492.1; -;
DR InterPro: IPR003593; AAA.
DR InterPro: IPR001140; ABC_transporter_tmam.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam: PF00664; ABC_membrane.1.
DR Pfam: PF00005; ABC_tran.1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
DR ATP-binding; Transport.
SQ SEQUENCE 580 AA; 61803 MW; 3A6B3696E3798B3C CRC64;

Query Match 42.6%; Score 1195; DB 2; Length 580;
Best Local Similarity 45.5%; Pred. No. 9.3e-65;
Matches 240; Conservative 106; Mismatches 176; Indels 6; Gaps 3;

OY 28 VFEFFINILALVSPLYMLOYDRVLTSRNVSTLIVTCVFLFLIVYGLLEALRTQVLYR 87
Db 24 LFSAVINLMLALVALVLMLOYDRVLTASGNPMTLLMTLMLGLFGLMGALEMWRSQVYIR 83
OY 88 GGLKFDGVARDPIFKSVLDSTLSRKSG--IGQAFRMDQVREFMTG-GLIACFADFWTP 143
Db 84 LGTQMDRLNQRVYDAFAEQL--KGTQAAGQALSDLTLLRQFATQALFAFEDAWFP 141
OY 144 VFIYVSMNLHPFEGILAIACIIIFGLAVNMNDATKNPIOMATYASIAQNDAGSTLRNA 203
Db 142 VYLLVLFLEFHPWLGVLALAGALMLMALVWNOHICQAPLVAQNLSSISAOATANLRNA 201
OY 204 EVNKAQMGWGLOARWRARREQVAMQAAASDAGAVMSGIKYFRNIVOTLILGGAYLA 263
Db 202 EALEAMQMLATLARKFAQOALFLAQNALASEKTAIASMSKVRALATGSLVGLGALLA 261
OY 264 IDKISAGAMIAGSILVGRALADIEGAVGQKNYIGARGMDRLQTMREEKADHMLPL 323
Db 262 VQCAITPGMMIAGSILMGVLSPLDILGVWKQWSSARLAYERLITMLANPARTERMSL 321

QY 324 PEEGVLSAEASILPEGAOQPTMRQASFRIDAGAAVALVGPSSAGKSSLLRGIVGWPIC 383
 DB 322 PARSQGLTEVQVACAPSSRRPALANLGFSLPAGDVLVGGPSCGKSTLRLRLVGMTP 381
 QY 384 AAGVITLDGDIKQNPBEKIGRHVGLIPQDIEFSGIVANIRFTEFEGEYIEAATLA 443
 DB 382 MAKRVRLDGDALQWQKQOQGGPHIGYLPQDIEFSGIVANIRFTEFEGEYIEAATLA 441
 QY 444 GVEHMIOSLPMGYDTAIEGEGASLGGQRORLALARVAFMPLLVDEPNASLDQVGEV 503
 DB 442 GYVQLILQLPQGDITQLEGAGLSGGQKORVALARLYLPLALIVLDEPNASLDQVGEV 501
 QY 504 ALMEAMKRLKAARKVYIFATHKVLLAQAADYIMVINGVISDFEGRDR 551
 DB 502 ALHAIITLKAORRTVLTHTKTPVLTALTDQLILTRDQGLQAFGPAPAR 549

RESULT 5
 Q92N19 PRELIMINARY; PRT; 583 AA.
 AC 092N19;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE ABC PROTEIN.
 GN APRDPF33.
 OS Pseudomonas fluorescens.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=294;
 RN 11)
 RP SEQUENCE FROM N.A.
 RC STRAIN-NO.33;
 MEDLINE-99453738; PubMed-10524213;
 RA Kawai E., Idei A., Kimura H., Shimazaki K., Akatsuka H., Omori K.;
 RT "The ABC-exporter genes involved in the lipase secretion are clustered
 with the genes for lipase, alkaline protease, and serine protease
 homologues in Pseudomonas fluorescens no. 33.";
 RL Biochim. Biophys. Acta 1446:377-382(1999).
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 DR EMBL: AB015053; BAA36463.1; -.
 DR MEROPS: C39, UPF003593; AAA.
 DR InterPro: IPR003593; AAA.
 DR InterPro: IPR001140; ABC_transporter_tmem.
 DR InterPro: IPR003439; ABC_transportr.
 DR InterPro: IPR001687; ATP_GTP_A.
 DR Pfam: PF00664; ABC_membrane; 1.
 DR Pfam: PF00005; ABC_tran; 1.
 DR SMART: SM00382; AAA; 1.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 DR ATP-binding; Transport.
 KW SEQUENCE 583 AA; 62919 MW; 153D27B628E189B8 CRC64;

Query Match 42.5%; Score 1193.5; DB 2; Length 583;
 Best Local Similarity 44.4%; Pred. No. 1.2e-64;
 Matches 244; Conservative 112; Mismatches 180; Indels 13; Gaps 3;

QY 1 MFRSGAKPTIFDQAVLVARPAVTAMVFSEFINILALVSPILMYQVDRVLTSSRNSTL 60
 DB 10 LFRKLSGYKST-----LISIGCFALINVLMLVPSILMYQVDRVLTSSRNSTL 58
 QY 61 IYLVAVCVFLFVLVGLLEALRTQVLRVGLKFDGVARDPTRKSVLDSTLSR-KSIGQOAF 119
 DB 59 VMELTLMVYGFPAFVLEIVRSFIVIRIGSLERFNLRYKAFERNLQSGEBHGQSL 118
 DB 119 GDLTHIQFITGPALFAFEDAPMPPIYLFVIFLNVMLGVLASAGAVLLIGLACINEYLT 178
 QY 120 RMDQVREPMFG-GLIAFCDAWPTPVFVIVSWMLHPFGLIATIIIFGLAVMNDAT 178
 DB 119 GDLTHIQFITGPALFAFEDAPMPPIYLFVIFLNVMLGVLASAGAVLLIGLACINEYLT 178
 QY 179 KNPITQMTASIAQNDAGSTLRNAEYKAKMGMGGLQARKRAREDEVYVQAASADAGG 238
 DB 179 KKPILGEASYSQOSTQTLATSHLHNAETIQAMGMLGALKRKWFKVHSOPLGLQNKASDTGS 238

QY 239 AVNSGKVERINIVQTLILGGAYLAIDKISAGAMTAGSLVGRALAPIEGAVGQKNYI 298
 DB 239 IISLSKSLRLCQSLVYLGLGALLVIKGDWTAGMTAGSLTMGRVLSPIQDLIAVWKMS 298
 QY 299 GARGANDRLQTMREKSSADDPHLPLEPRGVLSAEASILPEGAOQPTMRQASFRIDAGA 358
 DB 299 SARLAFORDLNMRPEPPPOEQMALPAPKGVNSFEQVSAQPPRRVPTLHQVSENLAGE 358
 QY 359 AVALVPSAAGKSSLLRGIVGWPICAGVIRLDGYDIKQNPBEKIGRHVGLIPQDIEFSG 418
 DB 359 VLGVLGASGSGKSTLRLVAVGWPITLGYRLDGCADHHRDRDLDLPHIGYLPQDIEFSG 418
 QY 419 GYVANIRFTEFEGEYIEAATLAGVHEMIOSLPMGYDTAIEGEGASLGGQRORLALA 478
 DB 419 GSADNIRAFRRDADPQLVVOAQAQGVHELILRLPHGYDTVLDEGSGLSGGQKORVALA 478
 QY 479 RAYFRMPALLYDEPNASLDQVGEVALMEAMKRLKAARKVYIFATHKVLLAQAADYIMV 538
 DB 479 RALYGGPRLVLDPEPNSTLDTVEGAALASAIMQKAGSTVLTVTHRSSALQADKLLVL 538
 QY 539 NOGVISDFG 547
 DB 539 NEGRLQAFG 547

RESULT 6
 Q92G94 PRELIMINARY; PRT; 578 AA.
 ID 092G94;
 AC 092G94;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE ABC TRANSPORTER TLID.
 GN TLID.
 OS Pseudomonas fluorescens.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=294;
 RN 11)
 RP SEQUENCE FROM N.A.
 RC STRAIN-SIK W1;
 RX MEDLINE-99173908; PubMed-10074078;
 RA Ann J.H., Pan J.G., Rhee J.S.;
 RT "Identification of the tlidEF ABC transporter specific for lipase in
 Pseudomonas fluorescens SIK W1.";
 RL J. Bacteriol. 181:1847-1852(1999).
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 DR EMBL: AF083061; AAD09853.1; -.
 DR InterPro: IPR003593; AAA.
 DR InterPro: IPR001140; ABC_transporter_tmem.
 DR InterPro: IPR003439; ABC_transportr.
 DR InterPro: IPR001687; ATP_GTP_A.
 DR Pfam: PF00664; ABC_membrane; 1.
 DR Pfam: PF00005; ABC_tran; 1.
 DR SMART: SM00382; AAA; 1.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 DR ATP-binding; Transport.
 KW SEQUENCE 578 AA; 62581 MW; 9EF2DED29432D697 CRC64;

Query Match 42.2%; Score 1184; DB 2; Length 578;
 Best Local Similarity 44.8%; Pred. No. 4.3e-64;
 Matches 243; Conservative 106; Mismatches 192; Indels 2; Gaps 2;

QY 12 PDQAVLVARPAVTAMVFSEFINILALVSPILMYQVDRVLTSSRNSTLIVTYICVLEF 71
 DB 20 FIOGGEKXSTLISVGCFTALINLMLVPSILMYQVDRVLTSSRNSTLVLTLVVGFF 79
 QY 72 LVYGLLEALRTQVLRVGLKFDGVARDPTRKSVLDSTLSR-KSIGQOAFRMDQVREMT 130
 DB 80 AFITGLEIVRSFIVIRIGSLERFNLRYKAFERNLQROGHAQALGDLTLRQFIT 139

| | | | |
|----------|---|--|--------------|
| OY | 131 | G-GLAFCDDAPWTEVEYVSMKLHPFGILLATIIACIIIFGLAVMNDKATKNPIOMATFMS | 189 |
| OY | 140 | GALFAPFDAPWEPFLYLIVLIFLFWNMLGVLATAGAVLLIGLACINEYLTKKPLGEAGAYS | 199 |
| OY | 190 | IAAONDAGSTLRNMEVYKAMGMGGLQARMRAREDEOVANQAAASDAGAVMSGIKVERN | 249 |
| DB | 200 | QOOSQSLATSHLHNMETIOAGMGLALKRKMFPVAVHSQFLGIONTASDGVITSLSKTRL | 259 |
| OY | 250 | IVOTLLIGGGAVLADOKISGAMIIAGSIIYGRALAIPEGAVGMKNTICARGAMDRLOT | 309 |
| DB | 260 | CLQSVLIGLGLALVYIKGDMTAKMMIAGSIIIMGRVLSPIIDLIAWKMKWSAKLAYORLDD | 319 |
| OY | 310 | MLREKSAADDMPIPEPPGVYLSAFAASIIPLPGAQOPIPMROASFRIDAGAVALVGPSSAG | 369 |
| DB | 320 | LLREPPDSEPMKPLPHPIGVSYSPFOVSAGPPGRRTPTLHOVSFTLGADEVGLVGLGASGSG | 379 |
| OY | 370 | KSSLRLIGVGVPCAAGYIRLDGYDIKOMPEKIGRHVGYLPDIDIEFSGTVAQNIARFT | 429 |
| DB | 380 | KSTLAVRLVGVWPTLGGTVRLDGCADIRHRMREDLGPPIIGLIPDIDIEFSSSIDNIARFR | 439 |
| OY | 430 | EESEGEVEATLGVHEHTMOSLPMGYDTAIEGEGASLSSGQORIALAARVFRMALLY | 489 |
| DB | 440 | QADPALVQAQQAQGVHEHLIRLPHGDTLLEGDGGGLSGGQORVALARALYGPRLIV | 499 |
| OY | 490 | LDEPNASLDQVEYVALMEAMKRLKAARVYIFATHKNNLQAQDIYINIQGYISDFGER | 549 |
| DB | 500 | LDEPNSLNDPVGGEALASAIYQMKQSSSVLYLTHRSSALAAQADKLLVNEGRLQRLARA | 559 |
| OY | 550 | DRC 552 | |
| DB | 560 | RRC 562 | |
| RESULT 7 | | | |
| | 054456 | | |
| ID | 054456 | PRELIMINARY; | PRT: 568 AA. |
| AC | 054456: | | |
| DT | 01-NOV-1996 (TREMBlrel. 01, Created) | | |
| DT | 01-NOV-1996 (TREMBlrel. 01, Last sequence update) | | |
| DT | 01-DEC-2001 (TREMBlrel. 19, Last annotation update) | | |
| DE | LiPB. | | |
| GN | LiPB. | | |
| OS | Serratia marcescens. | | |
| OC | Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; | | |
| OC | Serratia. | | |
| OX | NCBI_TaxID=615; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | STRAIN-SR41 | | |
| RX | MEDLINE=96062219; PubMed=7592412; | | |
| RA | Akatsuka H., Kawai E., Omori K., Shibatani T.; | | |
| RT | "The three genes liPB, liPC and liPD, involved in the extracellular | | |
| RT | secretion of the serratia marcescens lipase which lacks an N-terminal | | |
| RT | signal peptide."; | | |
| RL | J. Bacteriol. 177:6381-6389(1995). | | |
| CC | -I- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. | | |
| DR | EMBL; D49826; BAA08631.1; -. | | |
| DR | HSSP; P13569; INBD. | | |
| DR | InterPro; IPR003593; AAA. | | |
| DR | InterPro; IPR001140; ABC_transporter_tmem. | | |
| DR | InterPro; IPR003439; ABC_transporter. | | |
| DR | InterPro; IPR001687; ATP_GTP_A. | | |
| DR | Pfam; PF00664; ABC_membrane_1. | | |
| DR | Pfam; PF00005; ABC_tran; 1. | | |
| DR | SMART; SM00382; AAA_1 | | |
| DR | PROSITE; PS00211; ABC_TRANSPORTER; 1. | | |
| DR | ATP-binding; transport. | | |
| SO | SEQUENCE 568 AA; 64205 MW; 4FD37D4B30A0EC75 CRC64; | | |

| | | | | |
|-----------------------|-------------------|--------------------|-----------|-------------|
| Query Match | 42.1%; | Score 1182.5; | DB 2; | Length 588; |
| Best Local Similarity | 44.9%; | Pred. No. 5.4e-64; | | |
| Matches 236; | Conservative 110; | Mismatches 177; | Indels 3; | Gaps 3; |

[illegible]

| RESULT | 8 |
|--------|--|
| ID | 067993 |
| AC | 067993; |
| DT | 01-AUG-1998 (TREMBLrel. 07, Created) |
| DT | 01-AUG-1998 (TREMBLrel. 07, Last sequence update) |
| DT | 01-DEC-2001 (TREMBLrel. 19, Last annotation update) |
| DE | ZINC-PROTEASE TRANSPORTER. |
| GN | APRD. |
| OS | Pseudomonas fluorescens. |
| OC | Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae; |
| OC | Pseudomonas |
| OX | NCBI_TaxID:294; |
| RN | [1] |
| RP | SEQUENCE FROM N.A. |
| RC | STRAIN-CY091; |
| RC | MEDLINE=98162052; Pubmed=9501431; |
| RA | Liao C.H., McCallus D.E.; |
| RT | "Biochemical and genetic characterization of an extracellular protease |
| RT | from Pseudomonas fluorescens CY091."; |
| RL | Appl. Environ. Microbiol. 64:914-921(1998). |
| CC | -I- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. |
| DR | EMBL: AF004848; AAC38254.1; " |
| DR | HSSP: P13569; INBD. |
| DR | InterPro: IPR003593; AAA. |
| DR | InterPro: IPR001140; ABC_transporter_tmam. |
| DR | InterPro: IPR003439; ABC_transport. |
| DR | InterPro: IPR001687; ATP-GTP.A. |
| DR | Pfam: PF00664; ABC_membrane.1. |
| DR | Pfam: PF00005; ABC_tran.1. |
| DR | SMART: SM00382; AAA.1. |
| DR | PROSITE; PS00211; ABC_TRANSPORTER.1. |
| DR | ATP-binding; Protease; Transport. |
| SW | Sequence 579 AA; 62531 MW; 54052E01DDE9E0D7 CRC64; |

| | | | | |
|-----------------------|-------------------|--------------------|------------|-------------|
| Query Match | 41.9%; | Score 1177.5; | DB 2; | Length 579; |
| Best Local Similarity | 44.1%; | Pred. No. 1.1e-63; | | |
| Matches 242; | Conservative 113; | Mismatches 177; | Indels 17; | Gaps 5; |

| | | | |
|----|-----|---|-----|
| OY | 1 | MEKRGASPTTFEDQAVILVARRPAVITAAWFSFIIITLAVSLVMLQYDVRISRNSTL | 60 |
| | : | : : : : : : : : : : : : : | : |
| Db | 10 | LEKALGEKSL-----LISVCGFALINLMLNLPISYIMQYDVRILSSQNETTL | 58 |
| OY | 61 | IVLVIVCFELFVGLLEALPTQVLVNGKLFEDVANDPFIKSVLDSLIR-KIGCOAF | 119 |
| | : | : : : : : : : : : : : : | : |
| Db | 59 | VMLTLMVGFPAFIFGLEIVIRSFIVIRIGSOLERRFNRLRYKAAFERNLROGCOASTL | 118 |
| OY | 120 | RDMDVREFPMIG-GLIACDAPMTFVUYISWMLHPFGILAIICIIIFGLVAMNDAT | 178 |
| | : | : : : : : : : : : : : : | : |
| Db | 119 | GDUTIRHOFITGPALFAEFDAPEFITYELVFLEFNWGLVATAGALLILGLACINEYLT | 178 |
| OY | 179 | KNPIDMAIMASTIAONDSGTIRLNAEVMKAKMMNGLOAPRARARDEQVAAQAASAG | 238 |
| | : | : : : : : : : : : : : : | : |
| Db | 179 | KKPLCEA--SGFSOOSTOLAT--SAETIQAMGMIGALRQRFAYNAHDFLGONKASDTGS | 234 |
| OY | 239 | AVMSGIKYFERNIVQTLILGGGAYLADIKISAGAMIAGSILVGRALPITEGAVOMKNYI | 298 |
| | : | : : : : : : : : : : : : | : |
| Db | 235 | VITLSKSLRCLQSLGVILGLALVTYKCDMTAKMMIAGSILMGRLSPIDQLVAMKOWS | 294 |
| OY | 299 | GARGAMDRLQTMLEEKSSADHMLPDERGVLSAEALSIPLPGAQOPLMKRASRIDAGA | 355 |
| | : | : : : : : : : : : : : : | : |
| Db | 295 | SAKLIYORDELRLREFPEVEOMKLPARKGOVSFEQVSAAGPRGRMPTLHVSFNLAGE | 355 |
| OY | 359 | ANALVGSAPKSSILIRGVYWPCAACVIRIDGDIKCOMPERKIRGVYILPDIDIEFS | 418 |
| | : | : : : : : : : : : : : : | : |
| Db | 355 | VLGVLGASSGSKSTARLVAGWPLTAQTVIRIDGADHIRKBRDLDGPIGTLIPDIDIEFS | 418 |
| OY | 419 | GTYAONIRAFTEFSEOEYIEATTLAGVEMIQSLSPMGDYTAIGEGASLSGGOBORLATA | 475 |
| | : | : : : : : : : : : : : : | : |
| Db | 415 | GSIALDINARERADEERYQAQAQGVHEILIRLPHGYDTVLGDNGGSLSGGQORVALA | 474 |
| OY | 479 | RAVERPALVLIDEPNASIDQVGEVALMEAMKRLKAARKTYIIFATHKYNLLAQDYIMVI | 533 |
| | : | : : : : : : : : : : : : | : |
| Db | 475 | RALYGPRLIVIDEPNSNLDTFGEAALASATIQMAKQSSVLYLTHRSSALADQADKLVL | 534 |
| OY | 539 | NOGVISDFG 547 | |
| | : | : : : : : : : : : : : : | : |
| Db | 535 | NEGRLQAFG 543 | |

```

09XB64      RESULT      9
ID           09XB64      PRELIMINARY;      PRT;      588 AA.
AC           09XB64;
DT           01-NOV-1999 (TREMBLrel. 12, Created)
DT           01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT           01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE           PRD1 PROTEIN.
GN           PRD1.
OS           Erwina amylovora.
OC           Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC           Erwina.
RX           NCBI_TaxID=552;
RN           [1]
RP           SEQUENCE FROM N.A.
RC           STRAIN=eg;
RC           MEDLINE=99303693; PubMed=10373365;
RA           Zhang Y., Bak D.D., Heid H., Geider K.;
RT           "Molecular characterization of a protease secreted by Erwina
RT           amylovora.";
RL           J. Mol. Biol. 289:1239-1251(1999).
CC           -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR           EMBL: Y19002; CA442874.1; -.
DR           HSSP: P13569; INBD.
DR           InterPro: IPR003593; AAA.
DR           InterPro: IPR001140; ABC_transporter_tmem.
DR           InterPro: IPR003439; ABC_transportr.

```

DR InterPro: IPR001687; ATP_GTP_A-
DR Pfam: PF00664; ABC_membrane_1.
DR Pfam: PF00005; ABC_tran_1.
DR SMART: SM00382; AAA_1.
DR PROSITE, PS00211; ABC_TRANSPORTER_1.
KW ATP-binding; Transp.
SQ SEQUENCE 368 AA, 64317 MW, F8790B18470026B0 CRC64

| | | | | | | |
|-----------------------|-------|--------------|-------------|------------|--------|--------|
| Query Match | 41.9% | Score | 1176.5 | DB 2 | Length | 588 |
| Best Local Similarity | 44.7% | Pred | No. 1.3e-63 | | | |
| Matches | 234 | Conservative | 112 | Mismatches | 174 | Indels |
| | | | | | | 3 |
| | | | | | | Gaps |
| | | | | | | 3 |

| | | | |
|----|-----|--|-----|
| QY | 28 | VSEFSEINIALIVASPCYMLQÖVYDRAVLTAKSNVSLVILVIVLVEFVFLVYGVLGDEAKRTVLVK | 6 |
| Db | 26 | IFFAFINILMLVPSSVYMLQÖVDRVLPNSNETITLMLILMIGFVGMSLLEIVRSWIVIR | 85 |
| QY | 88 | GGKLRDGYAARDPIKFSVLDSTLSR- KGIIGCAFPMQDMQVREPMFG- GLIFCQAPMPMPF | 145 |
| Db | 86 | ISQSDMLRNLNRYVYTAALAEVNLKMGSSDQGLMDLITLROFLTGNALAFEDAPMPPIY | 145 |
| QY | 146 | VIVSWMLHPFEFGILAIACIILFGLAVMNDATKNPJOMATWMSIAQNDAGSTLNRAEV | 205 |
| Db | 146 | LMIIFLFWMGFLVGLVAGALLINELINELIASKPKPLBAKSKSLISGNSASTNLNRAEV | 205 |
| QY | 206 | MKAMGMGGLCARMPARDEOVAAQMAASDGGAVMSGIKYFRNIVOTLLILGGATLAD | 265 |
| Db | 206 | IALMLPLMKRMPSTLHQREFLNSORLASERAARTSTVRVRSNLSOVLVLGGGWLAD | 265 |
| QY | 266 | GKISAGAMIASIILGRALPIEGGAVGOMKVIYARGAMBLQTMREEKSADHMLPE | 325 |
| Db | 266 | GHMFGOMMIAGSILMGRLAPLIEQYITVWKSYSAAHSYGRVGLDPTHPQRTGMSLPR | 325 |
| QY | 326 | PRGVLSAERASTLPRG-AQOPTMROASPRIDAGAAVALVPSAAGKSSLRGLYGVNPCA | 384 |
| Db | 326 | PEGVIVSEBETPTPGSAGDAVLNHWVSIAIOPGEVLGIGSRSGKSTLRLVLGWIPVS | 385 |
| QY | 385 | AGVIRLDGDIKOMPBERKGRHVGYLPÖDIELESGTVANQNTARPEEESQVETEAATLAC | 444 |
| Db | 386 | EGIVRLDNDADITYOMKKDELGPYIGTLPDIDELFAGTIAENRNLDESEKVEIEMARLAC | 445 |
| QY | 445 | VHEMTIÖSLPMGYDTALIEGGASLVSOGOROLALARAVFRMFAALLVDEPNASLDQVEVA | 504 |
| Db | 446 | VHEMTIRPEKGYDSVLGNGAGSLSGGOKORIGLARALGDPSSLVILDEPNANLIDAGEKA | 505 |
| QY | 505 | LMAMKRLAARQVYTFANHKVNLDAQDYLIMVINOGLISPG | 547 |
| Db | 506 | LMDALILKÖRKNKTYLLIHRHLLISMTTRELILLVNGVNAFEG | 548 |

```

RESULT 10
.087808
ID 087808 PRELIMINARY; PRT: 583 AA.
AC 087808;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE EPRD PROTEIN.
GN EPRD.
OS Pseudomonas tolaasii.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=29442;
RN [1]
RP SEQUENCE FROM N.A.
RC SPRAIN=1116S;
RA Han B., Johnstone K.;
RT "Characterisation of a gene cluster required for synthesis and
RT secretion of extracellular metalloprotease by Pseudomonas tolaasii."
RL Submitted (JUL-1998) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL: AJ007827; CAAT07699.1; -.
InterPro: IPR003593; AAA.

```

DR InterPro: IPR001140; ABC_transporter_tmem.
DR InterPro: IPR003439; ABC_transportr.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam: PF00664; ABC_membrane; 1.
DR Pfam: PF00005; ABC_tran; 1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
DR ATP-binding; Transport.
KW ATP-binding; Transport.
SEQUENCE 583 AA; 62732 MW; 2E07BC58EBDCAFB CRC64;
SQ
Query Match 41.88; Score 1172; DB 2; Length 583;
Best Local Similarity 45.28; Pred. No. 2.3e-63;
Matches 238; Conservative 106; Mismatches 181; Indels 2; Gaps 2;
QY 23 VITAMVSEFFINILALVSPILMLOYDRLVSRNSTLIYIVLCVFLFYGLLEALRT 82
DB 21 LISVGCFTALINVLMPSTIMLOYDRLVSSQNETTLMVLYGFEAFIGTLEVRS 80
QY 83 QVLVRGGLKFDGVARDPIFKSVLDSTL-SRKIGGQAFRDQVREPMTG-GLIAFCDA 140
DB 81 FIVIRISQQLERRNRLNYKKAFFENLKSQGHAGQSLGDLTHIRQPLTGALFAFPDAP 140
QY 141 WTPVAVIVSMNLHPFGILAIITACIIIFGLAVMNDATKNPIQATMASIAQNDAGSTL 200
DB 141 WFPYVLEFIEFLNVLGVLATAGAVILLIALLACLINEYMTKKPLGEASGFSQSTOLATSHL 200
QY 201 RNAVVMKMGMGGLQARWRARRDOQVAMQAAADAGAVNSGIKVPFNIVOTLILGGGA 260
DB 201 HNAETIQMGMLGALKRKWFVNSQFGLQKMSADTGSVITSLSKSLNLCQSLVTLGGA 260
QY 261 YLAIDKISAGAMTAGSILVGRALAPIEGAVGQMKNYIGANGARDLQTMREEKADH 320
DB 261 LLVIGKEMTAGMILTAGSILMGKRVLSPIQDLIAVKKMSAKLAIQRIDELLEPREPEAQ 320
QY 321 MPLEPFGVISAELASILLPGAQOPTYMRQASFRIDAGANAVALVGPSPAKSSLLIGYGV 380
DB 321 MALPAPRGQVSFEQVNSAGPPERRMATLHOVSFHLGAGEVLGLVGLASGSKSTLARVLGV 380
QY 381 WPCAGVIRLDGYDIKOMDPKLGRIHGYLPDITELFSGTAONIAARTSESEVTEAA 440
DB 381 WPTLAGVIRLDGADITHMRDRLGPHIGYLPDITELFSGTADINIAARRDDPALVVOAA 440
QY 441 TLAVHEMIOSLPMGYDTAIGEGASLSGGQRRLALARAFAVRMPALVLDEPNASLDQV 500
DB 441 QQAQVHLLIRLPQGYDTVLGDNCGSLGQKQKVALARALYGAIRLIVDEPNASNDTV 500
QY 501 GEVALMEAMKRLKAARKTVYFAHKKVNLAAQADYIMVINOVSDFG 547
DB 501 GEALASAIYOMKKAQSSVVLVTHRSALAAQADKLAVLNEGRLQAFG 547
RESULT 11
Q9KGS6 PRELIMINARY; PRT; 599 AA.
AC Q9KGS6;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE APRD.
GN APRD.
OS Pseudomonas brassicacearum.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=86264;
RP [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEM421;
RX MEDLINE=21125577; Pubmed=11222613;
RA Chabeaud P., de Groot A., Bilter W., Tommassen J., Heulin T., Achouak W.;
RT "Phase-Variable Expression of an Operon Encoding Extracellular Alkaline Protease, a Serine Protease Homolog, and Lipase in

RT Pseudomonas brassicacearum";
RL J. Bacteriol. 183:2117-2120(2001).
CC -I- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL: AF286062; AAF87590.1; -
DR InterPro: IPR003593; AAA.
DR InterPro: IPR001140; ABC_transporter_tmem.
DR InterPro: IPR003439; ABC_transportr.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam: PF00664; ABC_membrane; 1.
DR Pfam: PF00005; ABC_tran; 1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
DR ATP-binding; Transport.
KW ATP-binding; Transport.
SEQUENCE 599 AA; 64291 MW; C2ADF5838CFB9D67 CRC64;
SQ
Query Match 41.58; Score 1164.5; DB 2; Length 599;
Best Local Similarity 43.68; Pred. No. 6.8e-63;
Matches 238; Conservative 121; Mismatches 184; Indels 3; Gaps 3;
QY 4 RSGAKPTIFDDAVLVARRAVITAMVSEFFINILALVSPILMLOYDRLVSRNSTLIYL 63
DB 5 RSVAAAPLE-KALGDYKNILISVGCFTALINVLMPSTIMLOYDRLVSSQNETTLM 63
QY 64 TVICVFLVYGLLEALRTQVLRGGLKFDGVARDPIFKSVLDSTLSR-KGIGQAFRDM 122
DB 64 SLMVVGFVFVIGLEMVNSFVIRISQQLERRNRLNYKKAFFENLKSQGHAGQSLGDL 123
QY 123 DVQREPMTG-GLIAFCDAWTPVAVIVSMNLHPFGILAIITACIIIFGLAVMNDATKNP 181
DB 124 THIRQPLTGALFAFPDAPMPPIYLLVLYLFPNVLGVFASIGTLLIQLACINEAMTKRA 183
QY 182 IQMTMASIAQNDAGSTLRAEAVEMKMGMGGLQARWRARRDOQVAMQAAADAGAVM 241
DB 184 LGQASVTSQOSTQLATSHLNAETIQMGMLGALKRSWFVHSRFLGLQNASDGAIVS 243
QY 242 SGIVERNIVOTLLIGGAVYLAIDKISAGAMISILVGRALAPIEGAVGQMKNYIGAR 301
DB 244 SLSTLRICLQSLVIGLALLVIGKEMTAGMILTAGSILMGKRVLSPIQDLIAVKKMSAK 303
QY 302 GAMPRLQTMREEKADHMLPFRGVLSAEASILLPGAQOPTYMRQASFRIDAGANA 361
DB 304 LAYRRLDALIQAFAPODDGMFLPAPRGQVSFEQVNSAGPPGGRNMTLQOVFSLSAAGEVL 363
QY 362 LVGPSAGKSSLLIGYGVWPCAGVIRLDGYDIKOMDPKLGRIHGYLPDITELFSGT 421
DB 364 VLGASGSKSTLARVLGVWPTFAGVIRLDGADITHMRDRLGPHIGYLPDITELFSGT 423
QY 422 AONTARFTEPESQEVTEATLAGVHEMIOSLPMGYDTAIGEGASLSGGQRRLALARA 481
DB 424 AENISRFQADPQKVVAQAQGVHELLIRLPQGYDTVLGEGGSLGQKQKVALARALAM 483
QY 482 FRMPALVLDEPNASLDQVGEVALMEAMKRLKAARKTVYFAHKKVNLAAQADYIMVINO 541
DB 484 YDRPSLVLDPEPNSLDVTGERALASAIAMKKTGTSVLVTHRSALAAQADKLVLDEG 543
QY 542 VTSDFG 547
DB 544 RLQAFG 549
RESULT 12
Q9HYJ8 PRELIMINARY; PRT; 600 AA.
AC Q9HYJ8;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE TRANSPORT PROTEIN HASD.
GN HASD OR PA3406.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.

OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PAOI;
 RX MEDLINE=20437392; PubMed=10984043;
 RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Gader R.L., Goltzy R., Tolenticino E., Westbrock-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kass A., Ladlig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PAOI, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 CC -I- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
 CC (ABC TRANSPORTERS).
 DR EMBL; AE004761; AAC06794.1; -.
 DR HSSP; P13569; INED.
 DR InterPro; IPR003593; AAA.
 DR InterPro; IPR001140; ABC_transporter_tmem.
 DR InterPro; IPR003439; ABC_transportr.
 DR InterPro; IPR001687; ATP_GTP_A.
 DR Pfam; PF00664; ABC_membrane; 1.
 DR Pfam; PF00005; ABC_tran; 1.
 DR SMART; SMO0382; AAA; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 DR ATP-binding; Complete proteome; Transport.
 SQ SEQUENCE 600 AA; 63454 MW; 33DE5F81A56FB2A CRC64;

| | | | | |
|-----------------------|-------------------|--------------------|-----------|-------------|
| Query Match | 41.28; | Score 1156; | DB 16; | Length 600; |
| Best Local Similarity | 43.48; | Pred. No. 2.3e-62; | | |
| Matches 230; | Conservative 117; | Mismatches 181; | Indels 2; | Gaps 2 |

| | | | |
|----|-----|---|-----|
| Qy | 20 | PRAYITAMVSEFFNIITLAVSPLMYLYDVLTSRKSVITLVTVCVFLVLYGLLEA | 79 |
| Db | 20 | RAIRSVALLTFVIMILMLAAISLYMLQYDDVLSGNNHMLTLMITLVLDTLLGALFW | 79 |
| Qy | 80 | LFTOVLVRSGLKDEVDANDPIFKSVLSTLSR -KGIGGAFRMDDOVREMTG -GLIAFC | 137 |
| Db | 80 | VASIVAVIRGGGLDQOLMORITYDASFNRSLERGEQAAGQALNDLTSLRQFTGNALPAFF | 139 |
| Qy | 138 | DAPMTPVFVYSMMLHPFEGILATITACIIITIGLAVMNDNATKNPDMATASIAAONDAG | 197 |
| Db | 140 | DAPWFFLYITLVTFESPMUGLITLALGALLVLVLWVNESRREPLAEGLSTILATQOAS | 199 |
| Qy | 198 | STLRNAEVMKAMWGGLQARWRARDEQVAMQAAASDAGAVNSGIKVRPNIVOTLLG | 257 |
| Db | 200 | ANLRQETLTAAGMLPAMRARWFAHQAFILARQNIGSERSAATGATSGVRLAOSLYLG | 259 |
| Qy | 258 | GGAYTLADCKISANGAMIASILVGLALAPFIEGAVGOMKNYTGAGANDRIOTMLREKSA | 317 |
| Db | 260 | LGAMTLAVDRLTPPGMIAIGSILMGVLSPLIDOLIAVWQMSGARQAQORLARILLEMPA | 319 |
| Qy | 318 | DHMLPEBRGVLSEAAISILPBGADQPTMRASPRIDAGAAVVLVSPGASGSSLRGI | 377 |
| Db | 320 | ALGMPILPAERGAIRVERILCAAAPGEEQALLDOLGFALEPGALGVITGSSGSKTLARL | 379 |
| Qy | 378 | VGWVPCAACAVILUDGYDIKOMPEKILGRHVGLIPDILIEFGTAONIAFRTEPESOEVI | 437 |
| Db | 380 | VGAMQPLSGAVRLDGADLRFQMSAALPHGHITGLADVOLFFGSLTAENIARAEVDAKVV | 439 |
| Qy | 438 | EAATLGAHVHEMIQSLPMGYDTAIGEGGASLSGGQRQRLALARAAYFRMDALLVDEPNASL | 497 |
| Db | 440 | AAARLGAHVHDVLRPEQGYDTRLGGGAGLSGGQRQRLGRLALRYGPRALIVLEDPNASL | 499 |
| Qy | 498 | DOYGEVALMEAKMKRLAAAKRTVIFATHKVNNLQAADYIMVINGVYSDFG | 547 |
| Db | 500 | DEGEALALEAIAMAROSSSVLVTYHPRAYLALTDKLLHLHGGSLDFEG | 549 |

| ID | PRELIMINARY; | PRT; | 582 AA. |
|-----|--|------|---------|
| AC | 054416; | | |
| DT | 01-NOV-1996 (TReMBLrel. 01, Created) | | |
| DT | 01-NOV-1996 (TReMBLrel. 01, Last sequence update) | | |
| DT | 01-DEC-2001 (TReMBLrel. 19, last annotation update) | | |
| DE | METALLOPROTEASE TRANSPORTER. | | |
| GN | HASD8000. | | |
| OS | Serratia marcescens. | | |
| OC | Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; | | |
| OX | NCBI_TaxID=615; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | STRAIN=SR1; | | |
| RA | Omori K.; | | |
| RT | "Cloning of the hasD8000 genes from Serratia marcescens SR1: | | |
| RT | secretion mechanism for the lipase and metalloprotease."; | | |
| RL | Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases. | | |
| CC | - i - SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. | | |
| DR | EMBL; D83582; BAA12015.1; -. | | |
| DR | HSSB; P13569; INBD. | | |
| DR | InterPro; IPR003593; AAA. | | |
| DR | InterPro; IPR001140; ABC_transporter_tmem. | | |
| DR | InterPro; IPR003439; ABC_transporter. | | |
| DR | InterPro; IPR001687; ATP_GTP_A. | | |
| DR | Pfam; PF00664; ABC_membrane; 1. | | |
| DR | Pfam; PF00005; ABC_tran; 1. | | |
| DR | SMART; SM00382; AAA; 1. | | |
| DR | PROSITE; PS00211; ABC_TRANSPORTER; 1. | | |
| KW | ATP-binding; Protease; Transport. | | |
| SEQ | SEQUENCE 582 AA; 62053 MW; 2FA93EF300F3689F CRC64; | | |

| | | | | |
|-----------------------|-------------------|--------------------|-----------|-------------|
| Query Match | 40.5%; | Score 1138; | DB 2; | Length 582; |
| Best Local Similarity | 44.2%; | Pred. No. 2.7e-61; | | |
| Matches 232; | Conservative 105; | Mismatches 186; | Indels 2; | Gaps 2. |

| | | | |
|----|-----|---|-----|
| Qy | 28 | VESFFINILIALVSPLYMLQVYTDVLTSSNNSTLLVTVICVFLPVYGLLEALRTQVYLR | 87 |
| Db | 16 | LETAAYINLMLPALPYMLQVYDVLPFSGNRRMTLMLTLVMVGYLFLMGLEIEMWRSQVYLR | 75 |
| Qy | 88 | GGLEKRDGYARDPFEKSVLDSLT-SKKIGGGAFFDMQOVREPMFG-GLIFCCAPMPFV | 145 |
| Db | 76 | LEAQMDDMLRNQVYAAEFETNKTGNPLAGALNDLTNLRFOFATGNALFPEFPAPEFVY | 135 |
| Qy | 146 | VIVSMVLHPFEFGLIALIACIIIFGLAVANDNATKNPIDMATMASIAQNDAGSTLNAEY | 205 |
| Db | 136 | LIVVELLHPWGLALASAGYIVYVLLAMLNQVRSQAPLAEGRALVLSMTQOANGLRAEA | 195 |
| Qy | 206 | KRMGMWGGLAARMRRARDEQVAAQAAASDAGAVMSGIKVFRNIYOTLLIGGAYTAID | 265 |
| Db | 196 | IAMGMLTDLRLRWLRHQHQFLLLQNRASEKIAATVATMSKTVRLATLSLIMGCALLAVN | 255 |
| Qy | 266 | GKISAGAMTAGSITLYGRALAPLEGAVGOMKNTIGARGMDLTQMLREKSSADHMLPE | 325 |
| Db | 256 | GDTTPGMMTAGSITLYGRVLPIDOLIGAMKQMSRSYQSLQSLREYVLAAANDPQAASLPLA | 315 |
| Qy | 326 | PRGVLSAAEASLILPFGAQOQPTWRQASFRIDGAAVLLGSPAAKSGSLILRGIVWPCAA | 385 |
| Db | 316 | PNSTLVNVLNOLASAPBGSAPVPHVGSFRLBCEGVYGVYGAAGSGKTLIMRLVYALTPLS | 375 |
| Qy | 386 | GVYRLDGYIDKQMDPEKLELRHVGYVPODIETFSGTVAONIAFTEFESQEVYEATLAVY | 445 |
| Db | 376 | GVYRLDGAIIIRQMDNQQLSPGLYVPODIQIFAGTUTENIAFQGVYAEKVVAAAALAVY | 435 |
| Qy | 446 | HEMIQSLPMGYDTAIEGGASLGGCORRLALRAVFMPLLVLYDEPNASLDQVEYAL | 505 |
| Db | 436 | HOIILHPLPGYTELEGGSGSLGGCORFVALARALYSGPALVYLDENANMIDREGGAL | 495 |
| Qy | 506 | MEAMKRLKAARKTVIFATHKVWLLAQAQYIWIYINGVSLDGERD | 550 |
| Db | 496 | LRIIEALKRGTIVLYVTHKPAILLTTDKLLVLAAGVQVHGHPSD | 540 |

RESULT 13
Q54416

•
•
•
•

RESULT 14
ID 092NX3 PRELIMINARY; PRT; 617 AA.
AC 092NX3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PUTATIVE TOXIN SECRETION ATP-BINDING ABC TRANSPORTER
DE PROTEIN.
GN SMC04207.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_Taxid=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21368234; PubMed=11474104;
RA Galibert F., Finan T.M., Long S.R., Puehler A., Abola P., Ampe F.,
RA Bariloy-Hubler F., Barnett M.J., Becker A., Boistard P., Bothe G.,
RA Boutry M., Bowser L., Buhrmester J., Cadieu E., Capela D., Chain P.,
RA Cowie A., Davis R.W., Dreano S., Federspiel N.A., Fisher R.F.,
RA Groux S., Godrie T., Goffeau A., Golding B., Gouzy J., Gurjal M.,
RA Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,
RA Kahn M.L., Kalman S., Keating D.H., Kiss E., Komp C., Lelaure V.,
RA Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B.,
RA Ramsberger U., Surzycki R., Thebault P., Vandenbol M.,
RA Vorhoefer F.J., Weidner S., Wells D.H., Wong K., Yen K.-C., Batut J.;
RT "The composite genome of the legume symbiont Sinorhizobium meliloti";
RL Science 293:668-672(2001).
DR EMBL; AL591789; CAC46614.1; -
KM Complete proteome.
SQ SEQUENCE 617 AA; 65748 MW; B7E513DEF98FD879 CRC64;

Query Match 38.8%; Score 1088.5; DB 16; Length 617;
Best local Similarity 39.4%; Pred. No. 2.9e-58;
Matches 220; Conservative 126; Mismatches 202; Indels 11; Gaps 6;
QY 1 MFRRSGKPIIFDOAVLV--ARRAVITAMVPSFFINITALVSPLYMQLVDRVTSRNV 58
DB 1 MFGNSKRTTPQLKAAAVIPAKRTAFGTGAIISGIVNLALTSPLFMQVYDVRVLSRLP 60
QY TLVLYVYICFELFVLVGLLEALRTQVLVGRGKFD---GVARPIKSVSDSTLSKGI 114
DB 61 TLVGLAVLATLMTGFMALVIVSRALRLRGESFDHRLSGVHAIVRPLETMM--PGD 118
QY 115 GGAQFRDMDQVREFMTG-GIARCDAPWTPVFLVSWMLHPFFGILAIICIIIFGLAVM 173
DB 119 GLQPLRLDINVGRFLAGAGTALFDLPWMPYLAIICFLPHFWIGMPLVGLVAVNMTLL 178
QY 174 NDNAITKPIQMATYASIAAONDAGSTLRNAEVKAMGMGGLARFARDEOVAMQAAA 233
DB 179 TDRLSVQPTDARFYGTDRNALMEARRNAETVRALGLERHLQORWLSANHLYLEARRA 238
QY 234 SDAGGAVMSGKIVERNVOTLLIGGAYVLAIDKISAGAMIASIIVGRALPIEGAVGQ 293
DB 239 ADAVAGAGSTSRMLRMILQSLILGAVMVLVEQVNSAGVMASSIMKGRALAPVDLAIGS 298
QY 294 WKNYTGARGAMDRLQTMIREKSSADDMPLPEPRGVLSAEASITLPGA--QOPTMRQASF 352
DB 299 WKPEFLMARQSMWDRRLNRLDILPRAHPRMPLPARKSILAVEGLTIVIPGSGSGKPIVAVGSF 358
QY 353 RIDAGAAVALVGSAAKSSILKGIQVWPCAGVIRLDGDIKQMPDEKIGRHVGLPQ 412
DB 359 SVPGGAGIIGIIPGSGSKSTTRALVGAMTPASGKVRLDGVSQOMRDEELGRIHIGLPQ 418
QY 413 DIELFSTVNONIATREFE-ESQEVIEATLVLAGVHEMIGQSLPMGYDAIGGSGSLSGGQ 471
DB 419 GVELFDSTIAENISRFNSNDPPTAIYAAQAGVHDLIVREQGYETRIGAGSALSAQ 478
QY 472 RQRLALAAVFRMPALLVLDEPNASLDQVEVALMEAMKRLKAKRTVIEATHRVNLAAQ 531

DB 479 RQRLGIALALYNDFLVLVDEPNANLDDAGKAVADILTSVRNREGIIVVAHBPALIGA 538
QY 532 ADLVMTVINOQVISPGEKD 550
DB 539 ADLVLVDSGGVNAFGPRD 557
RESULT 15
ID 033678 PRELIMINARY; PRT; 562 AA.
AC 033678;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PRSD (ABC TRANSPORTER ATP-BINDING PROTEIN).
GN PRSD.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_Taxid=382;
RN [1]
RP SEQUENCE FROM N.A.
RA York G.M., Walker G.C.;
RT "The Rhizobium meliloti exoK gene and prsd/prse/exsh genes encode components of independent degradative pathways which contribute to production of low-molecular-weight succinoglycans";
RL Mol. Microbiol. 25:117-134(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=1021; PLASMID=PSYMB (MEGAPLASMID 2);
RX MEDLINE=21396508; PubMed=11481431;
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
RA Vorhoefer F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
RA Golding B., Puehler A.;
RT "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-RT fixing endosymbiont Sinorhizobium meliloti";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR EMBL; U89163; AAB64091.1; -
DR EMBL; AL603646; CAC49692.1; -
DR InterPro: IPR003593; AAA.
DR InterPro: IPR001140; ABC_transporter_tmam.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam: PF00664; ABC_membrane_1.
DR Pfam: PF00005; ABC_tran; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_1.
KM ATP-binding; Plasmid; Complete proteome.
SQ SEQUENCE 582 AA; 62119 MW; 102FEF38C9ED6BA CRC64;

Query Match 38.4%; Score 1077; DB 16; Length 582;
Best local Similarity 42.2%; Pred. No. 1.3e-57;
Matches 226; Conservative 104; Mismatches 202; Indels 4; Gaps 3;
QY 20 RPAVITAMVPSFFINITALVSPLYMQLVDRVTSRNVSTIIVTVICVPLFYGLLEA 79
DB 19 RAAITGVGVSALVNLVLTGSSFFMLEVYDRILPSRSTPSIALSLALLAFARQAFEL 78
QY 80 LRTQVLVRGGLKFPGVARDPIFRKSVLDSTLS--RKGIGGAQFRDMDQVREFMTG-GIAR 136
DB 79 IRGMALYRIAGALDESUNGRIYRAIVAPLKLKNOGSGGLQALRPFDOYRSLSGVGRAL 138
QY 137 CDAPWTPVFLVSWMLHPFFGILAIICIIIFGLAVNNDNAITKPIQMATYASIAAONDA 196
DB 139 FDLFWLPEFYIAICFLPHVIGLAIIGGLITLTLYLTNRQTQAPARKASBAGGLRVFA 198
QY 197 GSTLRNNEVKAMGMGGLARFARDEOVAMQAAASDAGGAVMSGIVFRNIVOTLIL 256
DB 199 QASORNAEIVHAHGMASRLTALMERRNTEFRDENRRTSDINGYALSKYFRMALQSGVL 258
QY 257 GGGAYVLAIDKISAGAMIASIIVGRALAPIEGAVGQWKNYTGARGAMDRLQTMIREKS 316

```

Db 259 AAGAVLVIRGEASPGITIIAGSILTRALAPVELAIGNMRGLVARQSWORLKEELNALPE 318
QY 317 ADDHMLPEPRGYLSAEASILPPGAOQPTMRQASFRIDGAVALVGPSSAGKSSILRG 376
Db 319 ADAPLOLPDHERLTVEGIASGPPAQRVSDVNETVRAGAVGYIGPSAGKSSILARA 378
QY 377 IVGVWPCAGVIRLDGYDIKQMDPEKLRHVGYLPODIELFSGTVAQNIARFTE-FESQE 435
Db 379 ILGIWPAIRGVSRLDGDALDQWSDALGKHVGYLPQDVLEFAGTIAQNICRFADATSEA 438
QY 436 VIEAATLAGVHEMIQSLPMGYDTAIGBEGASLSGQORLALARAVERMPALLYLDEPNA 495
Db 439 IVAAKAKAARYNDLILRLPNGDYDEIGDGMTLISAGORVALARALYGDPFLVYLDEPNS 498
QY 496 SILDOVEVALMEAMKRLKAKKRTVIFATHKVNLLAQADYIMVINOGVYISDEGERDR 551
Db 499 NLDAEGQALSEAIMSVSRGCIYIVVAHRPSALASVDLYLMNENGRMQAFGEREQ 554

```

Search completed: July 19, 2002, 10:48:41
 Job time: 416 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2002, 10:43:17 ; Search time 55.96 Seconds

(Without alignments)
863.422 Million cell updates/sec

Title: US-09-913-414-5

Perfect score: 2165

Sequence: 1 MKPKIQRTDNFQAVARIG.....VLQYLESPLRDLRTTMRRE 435

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|--------------------|
| 1 | 2165 | 100.0 | 435 | 21 | AA822787 | Caulobacter cresce |
| 2 | 521.5 | 24.1 | 444 | 17 | AA805298 | Esterase secretory |
| 3 | 495 | 22.9 | 444 | 20 | AAV55920 | Pseudomonas fluore |
| 4 | 479.5 | 22.1 | 443 | 20 | AAV55923 | Pseudomonas fluore |
| 5 | 335 | 15.5 | 415 | 22 | AB852753 | Escherichia coli p |
| 6 | 329 | 15.2 | 475 | 21 | AAV75573 | Neisseria meningit |
| 7 | 313.5 | 14.5 | 477 | 18 | AAW22161 | Apixiid protein. |
| 8 | 313.5 | 14.5 | 477 | 21 | AAV51414 | A. pleuropneumonia |
| 9 | 262 | 12.1 | 478 | 18 | AAW22154 | Apixid protein. Ac |
| 10 | 262 | 12.1 | 478 | 21 | AAV51408 | A. pleuropneumonia |
| 11 | 261.5 | 12.1 | 435 | 21 | AAV75574 | Neisseria meningit |

| | | | | | | |
|----|-------|-----|------|----|----------|---------------------|
| 12 | 201.5 | 9.3 | 425 | 20 | AAV24069 | Salmonella typhimu |
| 13 | 149.5 | 6.9 | 543 | 22 | AA667295 | Amino acid sequenc |
| 14 | 142 | 6.6 | 355 | 22 | AA698387 | Escherichia coli p |
| 15 | 130 | 6.0 | 285 | 22 | AAU36320 | Pseudomonas aerugi |
| 16 | 128.5 | 5.9 | 1761 | 20 | AAV15457 | Human laminin beta |
| 17 | 126.5 | 5.8 | 412 | 21 | AAV90283 | N. meningitidis BA |
| 18 | 125.5 | 5.8 | 1839 | 21 | AAV94291 | Caenorhabditis ele |
| 19 | 124.5 | 5.8 | 659 | 22 | AA693061 | C glutamicum prote |
| 20 | 122.5 | 5.7 | 2779 | 22 | AB862371 | Drosophila melanog |
| 21 | 119 | 5.5 | 331 | 20 | AAV34473 | Porphyromonas ging |
| 22 | 119 | 5.5 | 365 | 20 | AAV34349 | Porphyromonas ging |
| 23 | 118.5 | 5.5 | 796 | 20 | AAW62798 | S. aureus SecA2 pr |
| 24 | 117.5 | 5.4 | 233 | 22 | AAW63878 | Human polypeptide, |
| 25 | 116 | 5.4 | 1154 | 22 | AAU32042 | Drosophila melanog |
| 26 | 115.5 | 5.3 | 1294 | 22 | AB863502 | Pseudomonas aerugi |
| 27 | 115.5 | 5.3 | 2472 | 22 | AAU33568 | Human huntingtin-1 |
| 28 | 115 | 5.3 | 1090 | 21 | AAV59270 | LexA/NuMA fusion p |
| 29 | 115 | 5.3 | 2192 | 18 | AAW21732 | GAL4/HA/NuMA fusio |
| 30 | 115 | 5.3 | 2272 | 18 | AAW21731 | Haemophilus influe |
| 31 | 114.5 | 5.3 | 390 | 22 | AAU35553 | Putative P. abyssal |
| 32 | 114.5 | 5.3 | 1177 | 22 | AA896721 | Novel human diagno |
| 33 | 114.5 | 5.3 | 1717 | 22 | ABG20672 | Novel human diagno |
| 34 | 114 | 5.3 | 1374 | 22 | AB869070 | Human male enhance |
| 35 | 113.5 | 5.2 | 399 | 21 | AAV81743 | Streptococcus pneu |
| 36 | 113.5 | 5.2 | 457 | 20 | AAW94867 | Accessory protein |
| 37 | 113.5 | 5.2 | 541 | 22 | ABG94625 | Human protein sequ |
| 38 | 113.5 | 5.2 | 1201 | 22 | ABG10512 | Novel human diagno |
| 39 | 113.5 | 5.2 | 1235 | 22 | ABG24756 | Novel human diagno |
| 40 | 113.5 | 5.2 | 1283 | 22 | ABG14964 | Novel human diagno |
| 41 | 113.5 | 5.2 | 1283 | 22 | ABG24943 | Novel human secret |
| 42 | 113 | 5.2 | 2117 | 22 | AAU32040 | Novel human diagno |
| 43 | 113 | 5.2 | 2400 | 22 | ABG30278 | Novel human diagno |
| 44 | 113 | 5.2 | 2415 | 22 | ABG30279 | Novel human diagno |
| 45 | 112.5 | 5.2 | 792 | 22 | AAU34228 | Staphylococcus aur |

ALIGNMENTS

RESULT 1
AAB22787 standard; protein: 435 AA.
ID AAB22787
XX AAB22787:
AC
XX 22-DEC-2000 (first entry)
DT
XX
DE Caulobacter crescentus membrane fusion protein (MFP), RseE.
XX
XX Membrane fusion protein; MFP; RseE; channel protein;
KW type I secretion system; C-terminal secretion signal;
KW Caulobacter host cell; biofilm formation; bioeffector;
KW recombinant protein expression host; sewage treatment;
KW wood pulping; fish vaccine.
XX
OS Caulobacter crescentus.
XX
XX WO200049163-A1.
XX
XX 24-AUG-2000.
XX
XX 21-FEB-2000; 2000WO-CAN0173.
XX
XX 19-FEB-1999; 99CA-2261186.
XX
XX (UYBR-) UNIV BRITISH COLUMBIA.
XX
XX SmIt J;
XX
XX WPI: 2000-571987/53.
XX
XX Caulobacter host cell for expression and secretion of a heterologous
XX polypeptide, useful for treating sewage, waste water and in the pulping
PT

PT process -
XX
PS Claim 1; Page 35; 46pp; English.

CC The invention relates to a Caulobacter host cell for expression and
CC secretion of a heterologous polypeptide. The Caulobacter host cell
CC comprises at least one surface layer transport protein having an
CC amino acid sequence homologous to the Caulobacter crescentus Rsad or
CC RsaE proteins (AA822786, AA822787). The host cell further comprises a
CC DNA construct comprising DNA encoding a polypeptide heterologous to a
CC surface layer (S-layer) protein of the cell, from and operably linked
CC to, a DNA encoding a Caulobacter surface layer protein secretion signal,
CC with the proviso that when the cell comprises transport proteins having
CC the same sequence as both the Rsad and RsaE proteins, the secretion
CC signal is not from the Caulobacter crescentus surface layer protein
CC (RsaE). The invention also encompasses a method for identifying a
CC Caulobacter suitable for use as a host cell for expression and secretion
CC of a heterologous polypeptide comprising the detection of homologues of
CC the rsad or rsae genes using a rsad/rsae hybridisation probe. The host
CC cell is used for the expression and secretion of a heterologous
CC polypeptide. The modified Caulobacter cells may be used to treat sewage
CC and waste water. The cells may also be grown in wood pulp suspensions and
CC then used in the wood pulping process. They can also be used in fish
CC vaccines. Caulobacter species are able to form biofilms, and can attach
CC themselves to surfaces without producing the extracellular enzymes or
CC polysaccharide slimes that are characteristic of most other
CC surface-attached bacteria. Caulobacter are therefore particularly suited
CC for use in bioreactor systems. The present sequence represents the
CC Caulobacter crescentus membrane fusion protein (MFP). RsaE. This protein
CC is anchored in the inner membrane of Caulobacter crescentus and is
CC thought to span the periplasm. Together with an outer membrane protein,
CC it forms a channel that extends from the cytoplasm through the two
CC membranes to the outside of the cell.

XX Sequence 435 AA;

Query Match 100.0%; Score 2165; DB 21; Length 435;
Best Local Similarity 100.0%; Pred. No. 2,6e-178;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKPPTIORTPTNFQAVNATIGIITLVGGLGNAAFPAPDSAVANGVSAEVSODVQH 60
DB 1 mkpptiortptnfqavnatigiiitlvvgllgnaaafpapsavangvsaevsqdvqh 60
QY 61 LEGGMLAKILVREGSEKAVAGVLELDPFOANAAGITRNCYVALKAMEARLAEEDQRP 120
DB 61 leggmilakilvregsekvagvylelptqanaaaqitrrngvalkamearllaeedqrp 120
QY 121 SISFPADLTQSARADPMVARAIADQAQFTERRQTIQGOVDLMNAQRLQYQSEIEGIDRQT 180
DB 121 sisfpadltqsaradpmvaraiadegaqfterrtqitgqvdlmnaqrlqyqseiegidrt 180
QY 181 OGKLUQGFIEDELIDKLKIDKGLVPRRLALPARAGSLSGISGRITADSKAVQAS 240
DB 181 ogklugfiedelidklklydglvprllalparagslsgritadskavqas 240
QY 241 DTQLVROIKQEFFEVSQSITETRVRLAEVTEKEVNASDAQRIKIVSPVNGTQONLRF 300
DB 241 dtqlvroi kqeffevsqsitetr vrlaevtekevasdaqr i k i v s p v n g t q o n l r f 300
QY 301 FTGAVNRAAEPLVDIADDEAFVIOAHFOPTDVNVIMGMWTEVRLPAFHSAGNDPDR 360
DB 301 ftegvnraaeplvdiaddeafviahfop t d v n v i m g m w t e v r l p a f h s a g n d p d r 360
QY 361 HDPAVAVADRISSPOKQARFLGLIVRVKQQLPRLRGRTAGMPQVYTPTEGRTYQL 420
DB 361 hdpavavadris spokqarf l g l i v r v k q q l p r l r g r t a g m p q v y t p t e g r t y q l 420
QY 421 FSPRLDRLRTMREE 435
DB 421 fsp r l d r l r t m r e e 435

RESULT 2

AAW05298
ID AAW05298 standard; Protein; 443 AA.

XX AAW05298;

DT 30-DEC-1996 (first entry)

DE Esterase secretory protein (ORF2 product).

KW Esterase secretory gene; protein secretion.

OS Serratia marcescens strain S141 (FERM BP-487).

PN EP733707-A2.

PD 25-SEP-1996.

XX 20-MAR-1996; 96EP-0104429.

PR 23-MAR-1995; 95JP-0063772.

PA (TANA) TANABE SEIYAKU CO.

PI Akatsuka H, Kawai E, Shibatani T;

XX WPI; 1996-444599/45.

DR N-PSDB; AAT39662.

PT New isolated esterase secretory gene from Serratia - used to

PS increase the productivity and extracellular secretion of esterases

CC Claim 3; Page 21-22; 29pp; English.

CC 3 Different proteins (AAW05297-99) participate in the mechanism of
CC secretion of esterase by Serratia marcescens S141. They are
CC encoded by an esterase secretory gene (AAT39662) that includes 3
CC open reading frames. The proteins can be produced by transformed
CC host cells, pref. S. marcescens or E. coli, carrying vector
CC plasmids incorporating the esterase secretory gene. The esterase
CC is secreted from the host cell, and is recovered from the culture
CC medium and from within the cells. The esterase is useful in
CC hydrolysis reactions.

XX Sequence 443 AA;

Query Match 24.1%; Score 521.5; DB 17; Length 443;
Best Local Similarity 29.3%; Pred. No. 1.5e-36;
Matches 133; Conservative 88; Mismatches 194; Indels 39; Gaps 7;

QY 6 IORPDNF-----QAVARIGYGIITLVGGLGNAAFPAPDSAVANGVSAEVS-Q 56
DB 5 ipepdnsyteeipgderrfttrngwlvvgllgflawaafapldkvaspsvtsvgnrk 64
QY 57 DVQHLGGMILAKILVREGSEKAVAGVLELDPFOANAAGITRNCYVALKAMEARLAE 116
DB 57 dvqhlggm ilakilvregsekvagvylelptqanaaaqitrrngvalkamearllae 116
QY 117 DQRPISFPADLTQSARADPMVARAIADQAQFTERRQTIQGOVDLMNAQRLQYQSEIEG 176
DB 117 dqrpisfpadltqsaradpmvaraiadegaqfterrtqitgqvdlmnaqrlqyqseiegi 176
QY 176 g1sivtisp1ldavkdkprvaei1a1qqlf1ast1q1seid-----gykxsmg1 177
DB 176 g1sivtisp1ldavkdkprvaei1a1qqlf1ast1q1seid-----gykxsmg1 177
QY 177 DRQTOGLKD-----QLGFIEDELIDKLKIDKGLVPRRLALPARAGSLSGISGRIT 229
DB 177 drqto glkd-----qlgfiedelidklklydglvprllalparagslsgrit 229
QY 229 grtq1g1q1g1s1r1nk1q1g1s1r1egm1smk1q1a1da1y1p1n1ry1ev1q1r1fa1evn1s1de1v 237
DB 229 grtq1g1q1g1s1r1nk1q1g1s1r1egm1smk1q1a1da1y1p1n1ry1ev1q1r1fa1evn1s1de1v 237
QY 230 ADRAKAVGASPTQKLRQIKQEFFEVSQSITETRVRLAEVTEKEVNASDAQRIKIVS 289
DB 230 adrakavgasptqk1rk1q1k1e1f1ev1s1q1s1i1t1et1rv1rl1a1ev1te1ke1v1na1s1da1q1r1i1k1i1v1s 289
QY 289 gr1g1q1k1q1l1es1q1r1d1q1r1f1a1d1y1q1rev1rt1q1a1q1md1ase1fr1nk1q1m1ad1a1gn1ta1ts 297
DB 289 gr1g1q1k1q1l1es1q1r1d1q1r1f1a1d1y1q1rev1rt1q1a1q1md1ase1fr1nk1q1m1ad1a1gn1ta1ts 297
QY 290 PVNGTRQNLRTFTGAVNRAAEPLVDIADDEAFVIOAHFOPTDVNVIMGMWTEVRLPA 349

Db 298 pvtgltvgnlftcggvgvgaagdhlmvpsqalivvdsrlkvqlfdkvygplvdimfta 357
 QY 350 FHAGNPDEPERHNPVAVADRID-----PQKARLFL---GIYRVVQKQPLRLRGRTA 401
 Db 358 fgnqtkpklpvtclvsadrlvdkangepyygmqvsvpegkmlsgedlkr----- 409
 QY 402 GMPAQIVPTGERTVLQYLFSPRLDRLRTTMRRE 435
 Db 410 gmpvefvtkgtrslslylflkpldrhcsltee 443

RESULT 3

AAY55920
 ID AAY55920 standard; Protein; 444 AA.

AAV55920;

15-FEB-2000 (first entry)

Pseudomonas fluorescens ABC transporter cassette component B.

ATP-binding cassette; transporter; operon; lipBCD; Serratia marcescens;

microbe; protein secretion.

Pseudomonas fluorescens.

JP11276172-A.

12-OCT-1999.

27-MAR-1998; 98JP-0080597.

27-MAR-1998; 98JP-0080597.

(TANA) TANABE SEIYAKU CO.

WPI: 1999-626936/54.

N-PSDB: AA222700.

An ABC transporter gene - used to increase the ability of a microbe to

secrete proteins

Claim 2; Page 16-17; 28pp; Japanese.

This sequence represents an ATP-binding cassette (ABC) transporter
 component B from an Pseudomonas fluorescens strain 33 and encoded by
 an ABC transporter operon. The operon comprises 3 genes where the
 termination codon of the first gene overlaps with the initiation codon
 of the second gene. The sequences of the encoded proteins have 60, 44
 and 46% homology respectively to the lipBCD proteins from Serratia
 marcescens. The novel gene and protein can give or increase the
 ability of a microbe for secreting a protein.

Sequence 444 AA;

Query Match 22.9%; Score 495; DB 20; Length 444;

Best Local Similarity 30.9%; Pred. No. 2.9e-34;

Matches 134; Conservative 86; Mismatches 201; Indels 12; Gaps 7;

9 PTFNFOAVNRIGIGIATLFWGLLG---WNAFAPLDSAVIANG-VVSAEVSQDVQHLGG 64

18 peltgarffarmg---lllvvgaggfllwaslapldqglpvggtlvvsgkrkavqlspg 74

65 MLAKILVRGEGKVKAGQVLFELDPTQANAAAGITRNQYVALKAMEARLLAERDQPSISF 124

75 vvarilvreegaevkgqplfrldqtngqdvshlgqymawasavawqserdngstltif 134

125 PADUTSQRADPMVAKALADEQAQFETEROTIQOVULMNAQRLOYSSEIGIDROTQGLK 184

135 peel-sgnpdaqalavlegqrqlfssrteafaregqitranlegataqngmrarsdlit 193

QY 185 DQGFTEDELDIDRLKLYDKGLVPRRLALEARAGSLSGSIGRLTRADRSKANQASDTOL 244
 Db 194 aqagslrdqlnmlqpladngylprrnlmeyerqlsqvgqdahtngesgvegllesrl 253
 QY 245 KVRQIKOEFEEOVSQSTTEFTRVRLAEYTEKEVVASDAQKIKIVSPVNGTAONLRFPTBG 304
 Db 254 klqhsaeeykvevrsladqqlrstleqqltsagfdlqhselnapadglavnlsvhteg 313
 QY 305 AVYRAAEPLVIDIAPDEEAFVIOAHFOPTVDVNVHMGVTEVRILPAFHSAGNPDEPHDPV 364
 Db 314 avrageclilelvpqdrlevegrlpvhlvdkvghlrvpallftalnqstrprvpgvsl 373
 QY 365 AVADRISDPQKQARLFLGIYRVVQKQPL-PHLGRV-TAGMPAQIVPTGERTVLQYLF 422
 Db 374 isadgmldektgmpry--vrltvssaleklnglvikpmpemeflrrgersllylflk 431
 QY 423 PLRDTLRTTMRRE 435
 Db 432 plldragsalte 444

RESULT 4

AAY55923
 ID AAY55923 standard; Protein; 443 AA.

AAV55923;

15-FEB-2000 (first entry)

Pseudomonas fluorescens ABC transporter cassette II component B.

ATP-binding cassette; transporter; operon; lipBCD; Serratia marcescens;

microbe; protein secretion.

Pseudomonas fluorescens.

JP11276172-A.

12-OCT-1999.

27-MAR-1998; 98JP-0080597.

27-MAR-1998; 98JP-0080597.

(TANA) TANABE SEIYAKU CO.

WPI: 1999-626936/54.

N-PSDB: AA222701.

An ABC transporter gene - used to increase the ability of a microbe to

secrete proteins

Claim 8; Page 23-24; 28pp; Japanese.

This sequence represents an ATP-binding cassette (ABC) transporter
 component B from an Pseudomonas fluorescens strain 33 and encoded by
 the ABC transporter operon II. The operon comprises 3 genes where the
 termination codon of the first gene overlaps with the initiation codon
 of the second gene. The sequences of the encoded proteins have 63, 56
 and 59% homology respectively to the lipBCD proteins from Serratia
 marcescens. The novel gene and protein can give or increase the
 ability of a microbe for secreting a protein.

Sequence 443 AA;

Query Match 22.1%; Score 479.5; DB 20; Length 443;

Best Local Similarity 29.0%; Pred. No. 6.3e-33;

Matches 130; Conservative 83; Mismatches 196; Indels 39; Gaps 7;

9 PTFNFOAV-----ARIGYGIATLFWGLGWAFAFAPLDSAVIANG-VVSAEVSQDVQHL 62

14 pvnvialddkkyarlgwllvlgfagflgwaalapltdkvavsgvkmvsgnhtkvtghps 73

```

QY 63 GGMJAKLIVREGCEVKKGOVFEELDPDQANMAAGITRNOYVALKAMERLAREDDRPSI 122
Db 74 g9ivierlevirgevsagqy1nhketrplvgmqmstqfisaasasariisaesegslv 133
QY 123 SFPADLTSGRADPVARAIADDEQAFTERROTIOGOYDLMNAORLOYOSETEGIDROTQC 182
Db 134 vfgeplln-----epaataatlsiqrglfssrsgala-----teggjireliagaeeqlrg 183
QY 183 LKD-----QLGTEDEBLIDKLKLYRKGLVPRRLLALPARAGSLSGSGRLTADRSKA 235
Db 184 tresqaskv1qrltameeq1qg1relatarday1prnrlleneriyag1dga1aedfgr1yq1 243
QY 236 VQASDNDLJKYROIKOEFEFEGVOSITRETRYRLAEVTEKEVNASDAQKRIKITYSPVNGTA 295
Db 244 qrvylelrlrlr1qgedqkql1yq1leetrssddllnr1laseefans1aratapaagv 303
QY 296 ONLRFTEGAVVRAEPLVJD1APDEDAFVIOAHNPVDVNVHMGWTEVRLPAFHSAGN 355
Db 304 vgl1dvyteggv1kpggqamdlvpbgger1lvearyrvpmvdkv1hpg1dvel1fsefngst 363
QY 356 PDPERHNDVAVALRISDFQOKARLF-----LGIYRVVQKOLPRHLGRVYTAGMAAQY 407
Db 364 pr1agevcl1vadrdqvdertdepyr1l1raqgsaaqm1qdg1qlr1-----gmprvea 415
QY 408 IVPGEERTVLOYLFSPRLDLTKRTMBRE 435
Db 416 fvk1etgersml1y1kpl1drthma1vee 443

```

RESULT 5
 ABB52753
 ID ABB52753 standard; Protein; 415 AA.
 AC
 XX ABB52753;
 DT 11-FEB-2002 (first entry)
 XX
 DE Escherichia coli polypeptide SEQ ID NO 907.
 XX
 KW Escherichia coli; B2/D+A-; antiinflammatory; antibacterial;
 KW immunosuppressive; extra-intestinal infection; phylogeny; meningitis;
 KW systemic infection; non-diarrhoeal infection; septicaemia;
 KW pylonephritis; antibiotic resistance.
 XX
 OS Escherichia coli.
 XX
 PN W0200166572-A2.
 XX
 PD 13-SEP-2001.
 XX
 PF 12-MAR-2001; 2001WO-EP03445.
 XX
 PR 10-MAR-2000; 2000FR-0003145.
 PR 02-FEB-2001; 2001FR-0001449.
 XX
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX
 PI Birgen E, Bonacorsi S, Clermont O, Nassif X, Tinsley C;
 DR WPI; 2001-550253/61.
 XX
 PT A library of DNA fragments of Escherichia coli strains for the
 PT phylogenetic determination of a given strain comprises polynucleotides of
 PT nature B2/D+ A- -
 PT
 XX
 PS Example 6; Fig 6; 646pp; English.
 XX
 CC The invention relates to a library of DNA fragments of Escherichia coli
 CC strains comprising polynucleotides (ABA88577-ABAB8729 and ABA89533)
 CC and encoded proteins (ABB52459-ABB52913 and ABB52954-ABB53094) of nature
 CC B2/D+A-. The polynucleotides have potential antiinflammatory,
 CC antibacterial and immunosuppressive activity as part of pharmaceutical

compositions used to treat, palliate or prevent extra-intestinal *E. coli* infections. The polypeptides are useful for determining the phylogenic group of a given *E. coli* strain. These polypeptides can detect and treat an undesired development of *E. coli*, particularly an extra-intestinal infection that include systemic and non-diarrhoeal infections such as septicaemia, pyelonephritis and meningitis this is particularly advantageous as bacterial resistance is increasing with the more frequent use of broad spectrum antibiotics.

| | | | | |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match | 15.5%; | Score 335; | DB 22; | Length 415; |
| Best Local Similarity | 25.6%; | Pred. No. 1.7e-20; | | |
| Matches 110; | Conservative 74; | Mismatches 162; | Indels 84; | Gaps 11 |

| | | | |
|----|-----|--|-----|
| Oy | 23 | IIATFPGSLGMAAFAPLADSAVIANG--VSAEVSODYOHJEGGLAKILVREGKVRAGQ | 81 |
| Dd | 47 | vlvltv---lwamspidevtvggsilpgstrevgicldlpgllkclevegdiuekq | 103 |
| Oy | 82 | VLEFLDPQANAAGGIRFMQYVALKAMEARLLAERDORPSISFPADLTGORAPMVARAI | 141 |
| Dd | 104 | vlltldtrrsamrlresearvnnleavrarlraea-ysestlfpdv----- | 149 |
| Oy | 142 | ADEQAOFFERKQOTOGGYDLMNAORLOYOSIEBIDKOTOGKQOLFIEDELIDRLKLY | 201 |
| Dd | 150 | ---padlrestv-----yrlktelagsiaqlkaskalllkeklamtrpiv | 193 |
| Oy | 202 | DGLVPRPRRLALBARAGSLGSGIRLADRSKAVOGASDPTQLVROIKOEFPEQVQSQI | 261 |
| Dd | 194 | regamsevellrmrgqsael-----qlmqdmekqmkylteagael | 233 |
| Oy | 262 | TEFVRVLAETVEKEVVASDAOKRIKIVSPVNGTAONLRFETEGAVVRAAEPLVDIADDE | 321 |
| Dd | 233 | vkteeelagakenagradpreersitcaprlgikvknlrvtlvgvsaagdimeilpded | 292 |
| Oy | 322 | AFVIOAHQOPFDVONVHNGMTEVRLPAFNSA-----GNPPERHDPAVADRI | 370 |
| Dd | 293 | qllleaylnpdrvagvyrtgmalvkltaeydalygldgvtlvtsptdltr-dqkrrgdlk | 351 |
| Oy | 371 | SDPOKARLFGIVRVVDKOLRPHLRGR-----VTAGMPAQVITVPGETVLOYLRSPL | 424 |
| Dd | 352 | ldpne-----ayrvlvtlsnnyltdtrngllpvlvpgmasvdlksgksvfygllkpl | 405 |
| Oy | 425 | RDTLRTTARE | 434 |
| Dd | 406 | -----trmkq | 410 |

| | RESULT | 6 |
|---------|--|----------------------------------|
| AA75573 | | |
| ID | AA75573 | standard; Protein; 475 AA. |
| XX | | |
| AC | AA75573; | |
| XX | | |
| DT | 21-MAR-2000 | (first entry) |
| XX | | |
| DE | Neisseria meningitidis ORF 764 | protein sequence SEQ ID NO:2620. |
| XX | | |
| KW | Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine; antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia; antibacterial; gene therapy. | |
| KW | | |
| XX | | |
| OS | Neisseria meningitidis. | |
| XX | | |
| PN | W09957280-A2. | |
| XX | | |
| PD | 11-NOV-1999. | |
| XX | | |
| PF | 30-APR-1999; | 99WO-US09346. |
| XX | | |
| PR | 01-MAY-1998; | 98US-0083758. |
| PR | 31-JUL-1998; | 98US-0094869. |

02-SEP-1998; 98US-0098994.
02-SEP-1998; 98US-0098062.
09-OCT-1998; 98US-0103749.
09-OCT-1998; 98US-0103794.
09-OCT-1998; 98US-0103796.
25-FEB-1999; 98US-0121528.
(CHIR) CHIRON CORP.
(GENO-) INST GENOMIC RES.
Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
Petersen J, Piza M, Rappoli R, Ratti G, Scalato E, Scarselli M;
Tetelhi H, Venter JC;
WPI; 2000-062150/05.
N-PSDB; AAZ54335.
Novel Neisserial polypeptides predicted to be useful antigens for
vaccines and diagnostics -
Claim 2; Page 1243; 1453pp; English.
AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ74253 to AAZ75941
represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides
and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
PCR primers used in the exemplification of the present invention. The
polypeptides, the polynucleotides, antibodies and compositions of
the invention can be used as vaccines, as diagnostic reagents, and as
immunogenic compositions. The polypeptides can be used in the
manufacture of medicaments for treating or preventing infection due to
Neisserial bacteria (e.g. meningitis and septicemia), to detect the
presence of *Neisseria bacteria*, or to raise antibodies. They may also
be used to screen for agonists or antagonists, which may themselves
have use as antibacterial agents. The polynucleotides of the invention
may also be used in gene therapy protocols.

[illegible]

Db 412 -heg---lglyfavwsldkhtlndgkavnltagmvtaelktgkrvldylspiqc 466

QY 427 TLRFTTMR 434
| : ||

Db 467 kidesfre 474

RESULT 7
 AAW22161
 ID AAW22161 standard; Protein; 477 AA.
 XX
 AC AAW22161;
 XX
 DT 16-FEB-1998 (first entry)
 XX
 DE ApxIIID protein.
 XX
 KW RTX toxin; apxICA gene; apxIIB gene; apxIIABCD gene;
 KW repeat in toxins toxin; cell-associated RTX toxin; vaccine production
 XX therapy; A. pleuropneumoniae infection; swine pleuropneumonia.
 OS Actinobacillus pleuropneumoniae.
 XX
 PN CA2170839-A.
 XX
 PD 02-SEP-1996.
 XX
 PF 01-MAR-1996; 96CA-2170839.
 XX
 PR 01-MAR-1995; 95US-0396244.
 XX
 PA (VXGU-) UNIV GUELPH.
 XX
 PI MacInnes J, Mallard B, Ricciatti P, Rosendal S;
 XX WPI; 1997-245536/23.
 DR N-PSDB; AAT73220.
 XX
 PT Preparations of microorganisms producing cell-associated RTX toxins
 XX pleuro-pneumonia
 XX
 XX Disclosure; Pages 113-114; 151pp; English.

CC AAM22151.7222161 represent A. pleuropneumoniae RTX (repeat in toxins)
CC toxins. These sequence are encoded by the apxICA, apxIIBD, apxIIAB'C, and
CC apxIIABCD genes (see AAT73211-73220), and can be expressed by
CC microorganisms used in the preparations of the invention. The
CC preparations are bacterial preparations comprising one or more isolated
CC and purified strains of a microorganism that produces one or more RTX
CC toxins, where the strains have at least one cell-associated RTX toxin.
CC The preparations are used for production of vaccines for the prophylaxis
CC and treatment of infectious diseases caused by microorganisms that
CC produce RTX toxins, where the strains have been attenuated or
CC inactivated. The vaccines are preferably against Actinobacillus
CC pleuropneumoniae infection (swine pleuropneumonia). It has been found
CC that A. pleuropneumonia produces significant quantities of
CC cell-associated RTX toxins when cultured under certain conditions, and
CC that the whole-cell protein composition of the cultures corresponds to
CC the whole-cell protein profiles obtained from cells recovered at
CC necropsy from the pleural fluid of infected swine. Vaccination with a
CC bacterin prepared from heat-inactivated cultures having significant
CC quantities of cell-associated RTX toxins give significant protection of
CC swine against challenge with homologous strains.
XX
XX Sequence 477 AA:

Query Match 14 5%: Score 313.5; DB 18; Length 477;
Best Local Similarity 25.3%; Pred. No. 1.5e-18;
Matches 112; Conservative 88; Mismatches 190; Indels 53; Gaps 13

```

Db      60  laylimifflaiv-1s1skveivaasatgklvfsgnskelkpiena1vdxlvykdgqfv 118
OY      78  KAGOVLEL-----DPTONNAAGITRNQYVALKAMEARLAEERQPSISF-PADLT 129
Db      119  ekqglllnlta1gdcadkqktkvs1glerldgyrks1ysi--ehnrplldngadfd 176
OY      130  S-GRADPMVARA1ADE-----OAFTERROTIGOVDMNAORLOYSEIEGIDR 178
Db      177  svgeedktgtrhllitegfetwqkqykelayqrkgaekqvtlanirkyesaarie---- 232
OY      179  QTOGLKQDLGFIEDLDLKLIVKGLVPRRLALAEAR---AGSISGSGRTADRSK 234
Db      233  -----keklsd1kkllygvks1skhellagenryveaanelsvyghlkeyesd 280
OY      235  AVOGASPTOLKROIKOEFPROVSQSTETPRRLAEVTEKEVNASDQKRK--IVSPVN 292
Db      281  l1kgedk1kvtqtkfksdllek1qgn1greqk1leleke-----qrqlas1lrpvs 334
OY      293  GTAQNLRFTEGAVVRAAEPLVDIAPDEDAFVIOAHFOPTDVNVHMGWTEVRLPAFHS 352
Db      335  gtvgqlkthkgyvvttaetlmv1apeddv1evsallqnkdvgtfvelgveavikvetfpy 394
OY      353  AGNDPDRHDPVAVADRI5DPQKARLFLGIVRVQKQLPRPHLG-RVTAQMPAQV1VPT 411
Db      395  lrygylgkvk1t1d1e1p1q-1g1vfns1le1nkk1t1d1gk1e1d1g1sgm1v1ae1kt 453
OY      412  GERTVLOVLFSP1RDLT1RT1T1MRE 434
Db      454  gersvis1fl1splees1tes1ire 476

```

```

RESULT 8
AAV51414
ID  AAV51414 standard; protein; 477 AA.
XX
AC  AAV51414;
XX
DT  05-MAY-2000 (first entry)
XX
DE  A. pleuropneumoniae apxIIID protein.
XX
KW  RTX toxin; vaccine; Repeats in Toxins; immunostimulatory; antimicrobial;
    antiinflammatory; antiaerobic; antiabortive; treatment; pneumonia;
    pleuropneumonia; septicemia; nephritis; arthritis; endocarditis;
    shipping fever; abortion; whooping cough; sleepy foal disease; joint ill;
    urinary infection; peritonitis; meningitis; gastroenteritis;
    passive immunization; apxIIID.
XX
OS  Actinobacillus pleuropneumoniae.
XX
PN  US6019984-A.
XX
PD  01-FEB-2000.
XX
PE  23-DEC-1996; 96US-0772270.
XX
PR  01-MAR-1995; 95US-0396244.
XX
PA  (UYGU-) UNIV GUELPH.
XX
PI  Mallard B, Rosendal S, MacInnes J, Ricciatelli P;
XX
DR  WPI: 2000-146864/13.
XX
DR  N-PSDB; AA288587.
XX
PT  Bacterial preparation comprising microorganisms which produce a member
    of the Repeats in Toxins (RTX) family, useful for treating swine
    pleuropneumonia, arthritis in swine, shipping fever and abortion in
    cattle, and sleepy foal disease -
XX
PS  Disclosure; Column 89-98; 96pp; English.
XX

```

```

CC      CC This invention describes a novel bacterial preparation (I) which
CC      CC comprises one or more isolated and purified strains(s) of a microorganism,
CC      CC cultured in tryptone yeast extract (TYE) broth, which produces one or
CC      CC more RTX toxins (belonging to the family of toxins referred to as Repeats
CC      CC in Toxins), where the strains(s) have at least one RTX toxin which is
CC      CC cell-associated. The products of the invention have immunostimulatory,
CC      CC antimicrobial, antiinflammatory, antiaerobic and antiabortive activity.
CC      CC The bacterial preparation may be used as vaccines for the prophylaxis and
CC      CC treatment of infectious diseases caused by strains of microorganisms
CC      CC which produce one or more RTX toxins. The infectious diseases are swine
CC      CC pleuropneumonia, pneumonia, septicemia, nephritis and arthritis in swine;
CC      CC septicemia, nephritis, endocarditis and arthritis in piglets; shipping
CC      CC fever and abortion in cattle; whooping cough, sleepy foal disease or
CC      CC joint ill (purulent nephritis, arthritis) in foals; septicemia,
CC      CC polyarthritis and abortion in horses; and urinary infections,
CC      CC peritonitis, meningitis, and gastroenteritis. The bacterial preparations
CC      CC may also be used to prepare antibodies which may be used as a means of
CC      CC passive immunization. This sequence represents the Actinobacillus
CC      CC pleuropneumoniae apxIIID protein described in the method of the
CC      CC invention.
SO      Sequence 477 AA;

```

```

Query Match 14.5%; Score 313.5; DB 21; Length 477;
Best Local Similarity 25.3%; Pred. No. 1.5e-18;
Matches 112; Conservative 88; Mismatches 190; Indels 53; Caps 13;

```

```

OY      19  IGYCIATFVGLGMAFAPLDSAVANG-VVSAEVSQDVQHEGMLKILVREGEKV 77
Db      60  laylimifflaiv-1s1skveivaasatgklvfsgnskelkpiena1vdxlvykdgqfv 118
OY      78  KAGOVLEL-----DPTONNAAGITRNQYVALKAMEARLAEERQPSISF-PADLT 129
Db      119  ekqglllnlta1gdcadkqktkvs1glerldgyrks1ysi--ehnrplldngadfd 176
OY      130  S-GRADPMVARA1ADE-----OAFTERROTIGOVDMNAORLOYSEIEGIDR 178
Db      177  svgeedktgtrhllitegfetwqkqykelayqrkgaekqvtlanirkyesaarie---- 232
OY      179  QTOGLKQDLGFIEDLDLKLIVKGLVPRRLALAEAR---AGSISGSGRTADRSK 234
Db      233  -----keklsd1kkllygvks1skhellagenryveaanelsvyghlkeyesd 280
OY      235  AVOGASPTOLKROIKOEFPROVSQSTETPRRLAEVTEKEVNASDQKRK--IVSPVN 292
Db      281  l1kgedk1kvtqtkfksdllek1qgn1greqk1leleke-----qrqlas1lrpvs 334
OY      293  GTAQNLRFTEGAVVRAAEPLVDIAPDEDAFVIOAHFOPTDVNVHMGWTEVRLPAFHS 352
Db      335  gtvgqlkthkgyvvttaetlmv1apeddv1evsallqnkdvgtfvelgveavikvetfpy 394
OY      353  AGNDPDRHDPVAVADRI5DPQKARLFLGIVRVQKQLPRPHLG-RVTAQMPAQV1VPT 411
Db      395  lrygylgkvk1t1d1e1p1q-1g1vfns1le1nkk1t1d1gk1e1d1g1sgm1v1ae1kt 453
OY      412  GERTVLOVLFSP1RDLT1RT1T1MRE 434
Db      454  gersvis1fl1splees1tes1ire 476

```

```

RESULT 9
AAW22154
ID  AAW22154 standard; Protein; 478 AA.
XX
AC  AAW22154;
XX
DT  16-FEB-1998 (first entry)
XX
DE  ApxID protein.
XX
KW  RTX toxin; apxICA gene; apxIIB gene; apxIIAB/C gene; apxIIABCD gene;
    repeat in toxins toxin; cell-associated RTX toxin; vaccine production;

```


| XX | Sequence | 478 AA: |
|----|---|---|
| XX | Query Match | 12.1%; Score 262; DB 21; Length 478; |
| XX | Best Local Similarity | 24.2%; Pred. No. 4e-14; |
| XX | Matches 111; Conservative | 86; Mismatches 178; Indels 84; Gaps 18; |
| OY | 19 IGYGIATLTFVCLLMAPAPLDSAVINGVVS-AEVSODVOHLESGMLAKLVREGKV 77 | |
| Db | 60 IAYLMLFLIALV-ISTVSHVELATATGKLAFSDRSKELPINALVKEIFVGDGCV 118 | |
| OY | 78 KAGGVLFELDPQANNAAGITRNQVVALKAMAR--LLAE---ROORPSISPAD--LTS 130 | |
| Db | 119 EKKDQILHHTALGADADGQTKSSISITKLERYEYELLVAAADRIPLEITKDEFNHA 178 | |
| OY | 131 GRADPMVARRAIDEAOQETER-----RDTIOGQVDLMNAORLOQSEIEGIDROT 180 | |
| Db | 179 TEEDKTRIRYLLTEGFQWQKQYKQKELAGREAEKQTVLANIK-----YEGISR-- 230 | |
| OY | 181 QGLKDLQGFIDE-LIDRLKLYDKGLVPRRLALEAR---AGSLSGSIGRLTDRSKA 235 | |
| Db | 231 -----VENERIKDKLKFNSKTSKSHDVLTEGNRIHEAVNELAYKSLNEVESQL 281 | |
| OY | 236 VQGASDQQLKVKNOIQEFPEQVQSQSTTERVRVLAETKEEVVASDAQRIKIV--SPVNG 293 | |
| Db | 282 RQAKEELHITLIFRADILEKIKQV--EAEKQISLEK-----NEQRJASVLTAPVSQ 335 | |
| OY | 294 TQONLRFEEGAVVRAAEPLVDIAPEDEAFVIAQHPQPDVDNVHMGVTEVRLPAFHSA 353 | |
| Db | 336 LVGQGLKTHVGVVTTAECLMVLAPEDDVLERYETALIGNKIDIGFIEVGQDAVKVELF--- 392 | |
| OY | 354 GNPDEPRHDPVA-----VADRISDPQKQARLFLGIVRVYDKQLPPLHRCR-- 398 | |
| Db | 393 -----PYRRYGLMGKVKNICTLEAHNPQ-IGLVFNLSISIDRKC---LSGKDGK 438 | |
| OY | 399 ---VTAGMPAQVIVPTGERTVLOQYLFSPRLDRLRTMRE 434 | |
| Db | 439 EELGSGMSVTAELKIGERSVISYILSPIEESVSELR 477 | |
| XX | RESULT 11 | |
| XX | AAV75574 | |
| XX | AAV75574 standard; Protein; 435 AA. | |
| XX | AAV75574; | |
| XX | 21-MAR-2000 (first entry) | |
| XX | Neisseria meningitidis ORF 764 protein sequence SEQ ID NO:2622. | |
| XX | Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine; | |
| XX | antigenic; diagnosis; immunogenic; infection; meningitis; septicemia; | |
| XX | antibacterial; gene therapy. | |
| XX | Neisseria meningitidis. | |
| XX | W09957280-A2. | |
| XX | 11-NOV-1999. | |
| XX | 30-APR-1999; 99W0-US09346. | |
| XX | 01-MAY-1998; 98US-0083758. | |
| XX | 31-JUL-1998; 98US-0094869. | |
| XX | 02-SEP-1998; 98US-0098994. | |
| XX | 02-SEP-1998; 98US-0099062. | |
| XX | 09-OCT-1998; 98US-0103749. | |
| XX | 09-OCT-1998; 98US-0103794. | |
| XX | 09-OCT-1998; 98US-0103796. | |
| XX | 25-FEB-1999; 99US-0121528. | |
| XX | (CHIR) CHIRON CORP. | |

PA (GENO-)INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizzo M, Rappuoli R, Ratti G, Scalati E, Scarselli M;
PI Tettelin H, Venter JC;
XX
DR WPI: 2000-062150/05.
XX
DR N-PSDB; AAZ54336.
XX
PT Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics -
XX
PS
XX
XX Claim 2: Page 1244; 1453pp; English.
XX
CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ53473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
CC presence of *Neisseria* bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
SQ Sequence 435 AA:

| Query Match | 12.1% | Score 261.5 | DB 21 | Length 435 |
|-----------------------|---|--|-----------|------------|
| Best Local Similarity | 25.2% | Pred. No. 3.9e-14 | | |
| Matches 102 | Conservative 70 | Mismatches 163 | Indels 69 | Gaps 14 |
| QY | 23 | IIATLTGVLGMAAFAPLDSAVIANG-VVSAEVSODVOHLESGMLAKILVREGKEVKAGQ | 81 | |
| DB | 62 | Imafallall- <u>ws</u> wfgkldivaasgkfvsggrsktlptletvvkavhvdqghvkqge | 120 | |
| QY | 82 | VLEFLDP-----QANNAAGITRNQOVA-LKAMEARLLARDDQPSISPADLTS | 130 | |
| DB | 121 | tlaelaavgtdsdvvsgegaigaqislkrlryeaavlalaestr----- <u>tl</u> | 163 | |
| QY | 131 | QRADPMVAFAIADDEQAQFTERRQTQGVDDLMNAORLOYQSEFESIDRQTGKLDQ---- | 186 | |
| DB | 164 | phidmagarslgsldadvsqavlaqhgygawaagaqqlsarlghqaelsakqeqk1 | 223 | |
| QY | 187 | --LGFIDE-LIDLKRLVDKGLVPRRLALAEARGS-----LSGSGITRUTDRSKAAVGA | 239 | |
| DB | 224 | vsvgaieqgktdyrrlradinfisehaflseqgsksvsmndlestrggmrqigaiaqae | 283 | |
| QY | 240 | SDTQDKRQIKOEFE--QVSQSITETFRVLAEVTEKEVVASDAQRIKIVSPVNGTAQ | 296 | |
| DB | 284 | qgrvntnqlkrctldalrtaqeqidqyrgq----tdk--akqyqlmtlqspadglvq | 336 | |
| QY | 297 | NKRFTGGVVAAPAVDIAPEDEAFVIOAHFQPTDVONVMGAVTEBRLPAF----- | 350 | |
| DB | 337 | elalcylvgvvgaaqkmmvavapddkmddevlvnkldigfegqgdaavklesfpyltyg | 396 | |
| QY | 351 | HSAGNPDEPRHDPVAVADRISDPQARLFLGIVREVDYKOLPPH | 394 | |
| DB | 397 | yltgkvksvshdavs-----heq-----lglyvtavvsldkh | 428 | |
| RESULT | 12 | | | |
| ID | AAV24069 | | | |
| XX | AAV24069 standard; peptide: 425 AA. | | | |
| AC | AAV24069; | | | |
| XX | | | | |
| DT | 09-SEP-1999 (first entry) | | | |
| XX | | | | |
| DE | Salmonella typhimurium peptide sequence SEQ ID NO:38. | | | |

```
XX Salmonella; Salmonellosis; detection; diagnosis; infection; food.
XX
XX Salmonella typhimurium.
OS
XX US5925522-A.
PN
XX 20-JUL-1999.
PD
XX 09-MAY-1997; 97US-0853659.
PF
XX 09-MAY-1997; 97US-0853659.
PR
XX 09-MAY-1997; 97US-0853659.
PA
XX (BATT ) BATTELLE MEMORIAL INST.
XX
XX Saifer JD, Wong K;
PI
XX WPI; 1999-418268/35.
DR
XX N-PSDB; AAX83950.
DR
XX
XX Detecting Salmonella in humans, animals and food
XX
XX
XX Disclosure; Column 161-164; 144pp; English.
PS
XX
XX The present sequence represents a Salmonella typhimurium peptide
CC sequence. The Salmonella nucleotide (AAX83935), its complement
CC (AAX83936) and fragments (AAX83937 to AAX83971), are useful for
CC detecting and diagnosing Salmonella infection in humans, animals and
CC food. The nucleotides will detect many, if not all Salmonella species
CC especially Salmonella dublin, S. enteritidis, S. gallinarum, S.
CC minnesota, S. paratyphi (types A, B and C), S. pullorum, S. typh ty21a
CC and S. typhimurium. The fragments may be used to construct a DNA chip,
CC useful for the simple and inexpensive testing of substances for the
CC presence of Salmonella.
CC
XX
XX Sequence 425 AA;
SQ
XX
Query Match 9.3%; Score 201.5; DB 20; Length 425;
Best Local Similarity 21.0%; Pred. No. 5.7e-09;
Matches 94; Conservative 93; Mismatches 187; Indels 73; Gaps 19;
QY 7 QRPDNEQAAVARIIGYITATLFEVGLGMAAFAPLDSAVIANGVSA-EVSQOVHLESGM 65
DB 3 rrgsdhlmml-----lsltl-l-llyfelnsvhggvlttkdnqllslskgvt 55
QY 66 LAKLIVREGKVKAGQV---FELDPQANAAGITRNQYVALKAMEARLLAERDQRPST 122
DB 56 lqdiylvaegdtvkkgeillakvnd-----lqkeyqyrtqkqyl--dkdvn-el 102
QY 123 SFPADLTSGRADPMV--ARAIADEQAQFERRQRTIOGOVULMNAQRLOQOSEIETGIDRQT 180
DB 103 sfldkenesgltlldgtrtslnkev--anlelvhsqitrlakelktsldeislqleki 160
QY 181 QGLKQGLFEDELIDIRKLYDKGLV-----RPRLLEARAGSLSGISGRFLTRDRS 233
DB 161 sskekelallaeenlnslprlykqispytnflnkqaylrvksesindessi-llkddi 219
QY 234 KAVOGASDTOLKVRQIOEFEFOVSQSITETRV--RLAEVTEKEVNASDAQRIKIVSPV 291
DB 220 elv--vndialmelnlrslsklslsknlqeleavnstclvlekkineed-----l 271
QY 292 NGRTQ--NLRFTEGAVVRAERPLVDIARPEDEAFVIAHQPTVDVNHMGMTVEVRLA 349
DB 272 dqvlyklnksactlmgvlgadallfelkprvtmldavkllprkyrdilvdeavklavqs 331
QY 350 FHSAGNDPDERHDPVAVADRISDPQO-----ARLEFLGIVRVDVQO-----LPRLRG 397
DB 332 l-----lqpkiksynatldnslpsyeentggtqrrykvvlaidvneddlrwlkp----- 382
QY 398 RVTAGMPAOVIVPTGERTVLOVLFSP 424
DB 383 ---gmctvdasvltgkhsimeyllspl 405
```

```
RESULT 13
AAG67295
ID AAG67295 standard; Protein: 543 AA.
XX
XX AAG67295;
AC
XX 13-NOV-2001 (first entry)
DT
XX
XX Amino acid sequence of ABC membrane fusion protein of avrXa21.
DE
XX
XX Avrulence gene; Xa21 resistance gene; avrXa21; type I secretion system;
KM pathogen infection; ABC membrane fusion protein.
KW
XX
XX Xanthomonas oryzae.
OS
XX
XX WO200162896-A2.
PN
XX
XX 30-AUG-2001.
PD
XX
XX 08-FEB-2001; 2001WO-US04068.
PF
XX
XX 22-FEB-2000; 2000US-0510206.
PR
XX
XX (REGC ) UNIV CALIFORNIA.
PA
XX
XX Ronald P, Shen Y, Da Silva FG;
PI
XX
XX WPI; 2001-514826/56.
DR
XX
XX N-PSDB; AAH77875.
DR
XX
XX Isolated polypeptide of an avirulence gene determinant specific for the
PT Xa21 resistance gene of plants is useful as a probe or ligand to
PT identify binding motifs and to develop a system to generate resistance
PT responses during infection -
PT
XX
XX Disclosure; Page 34-36; 47pp; English.
PS
XX
XX The present sequence represents an ABC membrane fusion protein. The open
CC reading frame is derived from a 9 kb avirulence gene determinant
CC specific for the Xa21 resistance gene from Xanthomonas oryzae pv.
CC oryzae, designated avrXa21. Several of the open reading frames of
CC avrXa21 encode predicted proteins with homology to the type I secretion
CC system of prokaryotes. AvrXa21 is secreted by avirulent Xanthomonas
CC oryzae pv. oryzae (Xoo) bacterial pathogens, and then interacts with
CC Xa21 gene products in the plant. The avrXa21 polynucleotide may be used
CC as a source of probes or ligands to identify the binding motifs of the
CC corresponding Xa21 R gene product. The avrXa21 polynucleotides and
CC polypeptides may also be useful in developing a two-component system to
CC generate resistance responses in plants in response to pathogen
CC infection.
CC
XX
XX Sequence 543 AA;
SQ
XX
Query Match 6.9%; Score 149.5; DB 22; Length 543;
Best Local Similarity 22.8%; Pred. No. 0.00024;
Matches 103; Conservative 57; Mismatches 172; Indels 119; Gaps 18;
QY 24 IALFFVGL--LGMAAFAPLDSAVI-ANGVSAEVSQOVHLESGMLKILVREGKVKAG 80
DB 171 vllllvgffclgfarrrtlygavvpadgmalttppq-----sgvanvgvvgqrvaaq 224
QY 81 QVLELDPQANAAGITRNQYVALKAMEARLLAERDQRPSTFPADLTSGRADPMVARA 140
DB 225 qvlflv-----aae-----hrrdr-----grpsqgaav 248
QY 141 IADQAQFERRQRTIOGOVULMNAQRLOQOSEIETGIDRQGLKQGLFEDELIDIRK- 199
DB 249 laeqg-----rltaeamvqlraggrlqqgaqa---ralaglnrlleyvdaelgvlfhr 298
QY 200 -----LVDKGLVPRRLALAEARAGSLSGISGRFLTRDRSKAVQASD 241
```

Db 299 qgtlqfiegryrtaltrglvsqgfvekqgadvidqrahal-----elqrerltladalaq 353
OY 242 TOLKVRQIKKEFEFQVS-----QSITETRVRLAEVTEKEVVASDAQRIKIVSPVNGTA 295
Db 354 agaelqglvpslltqqlalagaaslaqadrttaiegaasrwevrapragr----- 401.
OY 296 QNRFTEGAVVRAAEPLVDIAPEDEAFVIOAHFQPTVDNVHMGWTEVR---LPAFH- 351
Db 402 valrplqrggavqgqrldallpstatelvvlyapstraagllpglpvqtrfdalpyqhy 461
OY 352 -----SAGNDPERRHDPVAAADRIKDPOKQARFLGIVRVDVQQLPRHLRGVYTAG 402
Db 462 gqfagrvveiaaapepprad---aalaseplyrvrvla---gdaalragh-aavlrpg 513
OY 403 MPAQVIVTGERVLOYLFSPFLRDTLRTMKR 433
Db 514 mrvqglalalewrtisqwafepl-slingltlr 543
RESULT 14
AAC98387
ID AAC98387 standard; Protein: 355 AA.
XX AAC98387;
XX
DT 21-SEP-2001 (first entry)
XX
DE Escherichia coli protein sequence SRQ ID NO:435.
XX
KM Escherichia coli; identification: proliferation; microorganism:
KM antimicrobial; antibacterial; antibiotic; gene therapy; diagnosis;
KM bacterial growth inhibition.
XX
OS Escherichia coli.
XX
PN WC200148209-A2.
XX
PD 05-JUL-2001.
XX
PE 19-DEC-2000; 2000WO-US34419.
XX
PR 23-DEC-1999; 99US-0173005.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Forsyth RA, Ohlsen KL, Zyskind JW;
XX
DR WPI: 2001-457376/49.
DR N-PSDB: AAH81443.
XX
PT Novel nucleic acids encoding proteins required for Escherichia coli
PT proliferation, useful for screening for antimicrobial agents -
XX
PS Claim 19, Page 554-555: 596pp; English.
XX
CC The present invention describes a purified or isolated nucleic acid
CC sequence (I) consisting essentially of one of the 93 nucleotide sequences
CC given in AAH81292 to AAH81294, where expression of the nucleic acid in a
CC microorganism is capable of inhibiting proliferation of a microorganism.
CC (I) have antibacterial and antibiotic activities, and can be used in
CC gene therapy. Expression of (I) in a microorganism inhibits proliferation
CC of the microorganism, and the manufactured antibiotic is useful for
CC reducing the activity or level of a gene product required for
CC proliferation of a microorganism in a subject, specifically humans. The
CC nucleic acids that inhibit bacterial growth or proliferation can be used
CC as antisense therapeutics for killing bacteria. In addition to
CC therapeutic applications, the nucleic acid sequences complementary to
CC sequences required for proliferation can be used as diagnostic tools.
CC For example, nucleic acid probes complementary to proliferation-required
CC sequences that are specific for particular species of microorganisms can
CC be used as probes to identify particular microorganism species in
CC clinical specimens. AAH81295 to AAH81487 encode the Escherichia coli

CC proteins given in AAC98239 to AAC98431, and AAH81488 to AAH81491
CC represent oligonucleotides, which are used in the exemplification of the
CC present invention.
XX
XX Sequence 355 AA;
Query Match 6.6%; Score 142; DB 22; Length 355;
Best Local Similarity 21.2%; Pred. No. 0.0006;
Matches 87; Conservative 65; Mismatches 158; Indels 100; Gaps 20;
OY 22 GIIATLFVGLGNAAFAPL---DSAVIANGVSAEVSQDVHLEGMLKILVREGSKYK 78
Db 14 gllav--aaivawwllrpagypegfavngrlea-tevdlaaskiagrldilvkegkfvr 70
OY 79 AGCVLFEFLDPTQANNAAGITRNQYVALKAMEARLLAERDORPSISPADLTSORADBMVA 138
Db 71 egevlakmd-----tr-----viquegrleaaqikeags----- 99
OY 139 RAIVDEQAQFTEKRTIQGVVDLMNAQRLOYQSPFIEGI-DROTQGLK-DQLGFIEDELID 196
Db 100 -avaaagalllegqsetraagslvrn---grgaeldsvakhrtrsrjaqrqalsaqld 154
OY 197 LRLKYDKGLVPRPRLALLEARAGSLSGSIGRLTFDRSKAVOGASDTOLKVRQIKKEFEFQ 256
Db 155 -----dtraaesarralesakqyasakaieartnliq----- 190
OY 257 VSQSITETRVRLAEVTEKEVVASDAQRIKIVSPVNGTAQNLRFTGAVVRAAEPLVDI 316
Db 191 -----aglrvaagaqaterriaadiddselk--aprdgrvq-yrvaepegvlaagrvllm 242
OY 317 APEDEAFVIOAHFQPTD-VDNVHMG-----WTEVRLPA-----PHSAGNDPERRHDP 363
Db 243 v--disdymtflfippteagkligearllldaapdlrpatistfvaasqftrk---t 297
OY 364 VAVADRISDPQOKARFLGIVRVDVQQLPRHLRGVYTAGMPAQVYPTGE 413
Db 298 veld-----erlkimfvkarippellqghle-yvktglpgevawrvne 341
RESULT 15
AAU36320
ID AAU36320 standard; Protein: 285 AA.
XX
XX AAU36320;
XX
XX
DT 14-FEB-2002 (first entry)
XX
DE Pseudomonas aeruginosa cellular proliferation protein #310.
XX
KM Antisense: prokaryotic cellular proliferation protein;
KM antibiotic; antibacterial; drug design.
XX
OS Pseudomonas aeruginosa.
XX
XX WC200170955-A2.
XX
PN 27-SEP-2001.
XX
PD 21-MAR-2001; 2001WO-US09180.
XX
PE 21-MAR-2000; 2000US-191078P.
XX
PR 23-MAY-2000; 2000US-206848P.
XX
PR 26-MAY-2000; 2000US-207727P.
XX
PR 23-OCT-2000; 2000US-242578P.
XX
PR 27-NOV-2000; 2000US-253625P.
XX
PR 22-DEC-2000; 2000US-257931P.
XX
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;

XX MPI: 2001-611495/70.
DR N-PSDB; AAS54179.
XX
PT New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Example 3; Seq ID No 11913; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 285 AA:

Query Match 6.0%; Score 130; DB 22; Length 285;
Best Local Similarity 21.2%; Pred. No. 0.0048;
Matches 87; Conservative 56; Mismatches 103; Indels 164; Gaps 20;
QY 23 IIALTEVGLGMAFA---PLDSAVIANGVSAEVSQDVQHLEGMLAKILVREGKVK 78
DB 14 vaaavlaglwlwrymispwtrdarvadvv---vapdv---sgwvtdlevkdngyvk 66
QY 79 AGOVLEFLDPTQANMAAGITRNQYVALKAMEARLLAERDQPSISFPADLTQORADPVA 138
DB 67 vgdvllmrld-----geryqanleq-----a 86
QY 139 RAIADBEAQTERRTIQGOVDLMNAQRLQYOSIEIDRQTGLKQDLGFIEDLIDL 198
DB 87 ravaeltrh-----qy-----llrqneaar----- 106
QY 199 KLYDKGLVPRRLALEARAGSLSGISRLTADRSKAVOGASPTOLKVRQIKOEFPQVS 258
DB 107 -----tsrl-----qfgalsaecke-----naqinaalalseygeala 139
QY 259 QSTETFRVRLAEVTEKEVVASDAOKRIKIVSPVNGTAQNLR---FTEGAVVRAAEPL 313
DB 140 q-----vklaeln-----lkrseleaarngvtnrlagynyataqgamma---l 181
QY 314 VDIAPEDAEAFVIAHQPTVDVNHMGMTVEVRLPAFHSAGNPRPERHDPVA--VADRIS 371
DB 182 vd-----qgstfyvayfeetklpglirvgmaqvrl---msgdqpdlgtvesisgltids 234
QY 372 DPQKQ-----ARLEFLGI-VRVDVVKQLPRHLRGRTAGMPAOIV 409
DB 235 tpdgqlaanveptfnwrlqgtrlprlrlidqvpadv--hlsagmtasvty 282

Search completed: July 19, 2002, 10:43:19
Job time: 139 sec

THIS PAGE BLANK (USPTO)


```

Db 5 IGEPQDSTTEIEIPQDERFRFTMGVLVYGIGLFGFLAMAAAPADKVAAPSGVTVSGNRK 64
Oy 57 DVOHLEGMLAKILIVREGEKKVACQVLELDPPTQANAAAGITTRNOYALKAAMEARLLAER 116
Db 65 TVQVAPASQILKINIVRQDKKAAEVLVOLSVOYAQAVSLRQYVYTTLTATGRIILAE 124
Oy 117 DQRESIFSPADLTQSRADPWVARILADEQAQFERROTIGOVADLYMAARLOVQSEIEGI 176
Db 125 DGLSVTFTSPILDAVKDKPRVAEITAIQTLQFASRRALQSEID-----GKQSMDCI 177
Oy 177 DRQTOGLND-----QIGFIEDELIDLRLKYDKGLVPRPRLALEARAGSLSGSIGRLT 229
Db 178 RFOLKGLQDSGRNKQIQLSTSLREQMSNMKQLAAGLYPRNRXYLEVORQFLEVNSSIDEVY 237
Oy 230 ADRSKAVAGASDQTLKVAQIQOEFEFYSSQITTRVRLAEVTRKEVVAADACKRIKIVS 289
Db 238 GRIGLOQLQLESQORIDORFADYQREVRQTLAQTOQDASEFRNKLOMAOPDLGNATITS 297
Oy 290 PVNGTAQNLREFTEGAAVYRAALEPVLVDIAPEDEAVIAQHPPTQDVVNHMGMTVEVRLPA 349
Db 298 PVDGTIVGVGLNFTTQGGVAGAGDHLMDVVPQSATILVDSRLKYVDLFDKAVYNCLPVDLMEFA 357
Oy 350 FHSAGNPDPERRHDVAVADRISD-----PQKQARLFL-----GIVEVDVKQLEPHLRGHVTA 401
Db 358 FMOQNTPKPIGVTVLVSADRLVYDKRANGEPYQOMQVTVSPBGMKMLSGEDIKP----- 409
Oy 402 GMPRAQVIVPTGERIVLYQVLESPLADTLKTRTIRREE 435
Db 410 GMPVEVFWKTSRSLSTLYLEFPIIDRAHTLSTLEE 443

```

```

1      RESULT      2
2      US-08-620-605D-3
3      ; Sequence 3, Application US/08620605D
4      ; Patent No. 5846811
5      ;
6      ; GENERAL INFORMATION:
7      ; APPLICANT: SHIBAYANI, TAKEJI
8      ; APPLICANT: AKATSUKA, HIROYUKI
9      ; APPLICANT: KAWAI, ERI
10     ; TITLE OF INVENTION: GENE PARTICIPATING IN THE MECHANISM OF
11     ; TITLE OF INVENTION: SECRETION OF ESTERASE
12     ; NUMBER OF SEQUENCES: 4
13     ; CORRESPONDENCE ADDRESS:
14     ; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
15     ; STREET: PO BOX 747
16     ; CITY: FALLS CHURCH
17     ; STATE: VA
18     ; COUNTRY: USA
19     ; ZIP: 22040-0747
20     ;
21     ; COMPUTER READABLE FORM:
22     ;
23     ; MEDIUM TYPE: Floppy disk
24     ; COMPUTER: IBM PC compatible
25     ; OPERATING SYSTEM: PC-DOS/MS-DOS
26     ; SOFTWARE: PatentIn Release #1.0, Version #1.30
27     ; CURRENT APPLICATION DATA:
28     ; APPLICATION NUMBER: US/08/620,605D
29     ; FILING DATE: 22-MAR-1996
30     ; CLASSIFICATION: 435
31     ; ATTORNEY/AGENT INFORMATION:
32     ; NAME: SVENSSON, LEONARD R
33     ; REGISTRATION NUMBER: 30,330
34     ; TELECOMMUNICATION INFORMATION: 0020-3955
35     ; TELEPHONE: (703)-205-8050
36     ; TELEFAX: (703)-205-8050
37     ; INFORMATION FOR SEQ ID NO: 3:
38     ; SEQUENCE CHARACTERISTICS:
39     ; LENGTH: 443 amino acids
40     ; TYPE: amino acid
41     ; STRANDEDNESS: Single
42     ; TOPOLOGY: linear
43     ; MOLECULE TYPE: peptide
44     ;
45     ; US-08-620-605D-3

```

[illegible]

RESULT 3
; Sequence 13, Application US/08772270A
; Patent No. 601984
; GENERAL INFORMATION:
APPLICANT: Macinnes, Janet
APPLICANT: Ricciatti, Paul
APPLICANT: Mallard, Bonnie
APPLICANT: Rosendal, Soren
TITLE OF INVENTION: NOVEL BACTERIAL PREPARATIONS, METHOD FOR
NUMBER OF SEQUENCES: 14
PRODUCING SAME, AND THEIR USE AS VACCINES
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bereskin & Parr
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/772,270A
FILING DATE: December 23, 1996
CLASSIFICATION: 4/24
ATTORNEY/AGENT INFORMATION:
NAME: Gravelle, Micheline
REGISTRATION NUMBER: 40,261
REFERENCE/DOCKET NUMBER: 6580-81
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311


```

; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 477 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Actinobacillus pleuropneumoniae
US-08-772-270A-13

```

```

Query Match      14.5%; Score 313.5; DB 3; Length 477;
Best Local Similarity 25.3%; Pred. No. 1.1e-21;
Matches 112; Conservative 88; Mismatches 190; Indels 53; Gaps 13;

```

```

QY 19 IGYIATLTFVGLGMAAPPLDSAVIANG-VVSAEVSQVQHLEGMLAKILVREGKV 77
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 60 IAYLIMLFLELAV-LSISKEIVASATGKLVFSGHSEIKPIENALVKDIEVKDQFV 118
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 78 KAGOVLEL-----DPTQANAAAGTRNQYVALKAMEARLLAERDPSISF-PADLT 129
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 119 EKQILLNLTLALGCDADKQTKVSLGERLDGYRKSLSYI--EHNRLPLDPMQDFD 176
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 130 S-ORADPMVARAIAD-----QAQFERRQITQGVDMNAQRLQYSEIEGIDR 178
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 177 SVGEOKTGARHLITEGFETWQOKYKELAYRKQAKQTVLANIRKYSASIE----- 232
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 179 QYQGLKDLQGLFIEDELIDRLKLYDKGLVPRRLALAEAR---AGSISGSIGRLTADRSK 234
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 233 -----KEKLSLTKLYDVKSISKHELLAENRVEASNELSYQSHLKEVESD 280
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 235 AVOGASDTOLKVRQIKQEFEEYOSITETRVRLAEYTEKEVNASDQKRKIY-IVSPVN 292
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 281 LIAQEDLKVLTQLEFSDILEKLOONOREKQVLTLEKNE-----QROLASITRPVS 334
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 293 GTAONLREFTEGAVVRAEPLVDIAPEDEAFVIOAHFQPTVDVNMGMVTEVRLPAFHS 352
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 335 GTVQQLKTHVGGVVTAEFLMVAIPEDDVLEVSALQONKDVGFELGQDAVIVKVEFPY 394
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 353 AGNPDEPRHDVAVADRIADPQKARLFLGIVRVVYKQLPRHLRG-RVTAGMPAQVIVPT 411
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 395 TRYGYLGVKVTITLDAIEHPQ-LGLVFNLSIIEINKTTLTDGKEIOLGSGMSVIAEIKT 453
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 412 GERTVLOYLESPLDRLRTMRE 434
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 454 GERSVLSFLSPLSESTESLRE 476
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 4
US-08-772-270A-5
; Sequence 5, Application US/08772270A
; Patent No. 6019984
; GENERAL INFORMATION:
; APPLICANT: Macinnes, Janet
; APPLICANT: Ricciattl, Paul
; APPLICANT: Mallard, Bonnie
; APPLICANT: Rosendal, Soren
; TITLE OF INVENTION: NOVEL BACTERIAL PREPARATIONS, METHOD FOR
; TITLE OF INVENTION: PRODUCING SAME, AND THEIR USE AS VACCINES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bereskin & Parr
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/772,270A
; FILING DATE: December 23, 1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Gravelle, Micheline
; REGISTRATION NUMBER: 40,261
; REFERENCE/DOCKET NUMBER: 6580-81
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Actinobacillus pleuropneumoniae
US-08-772-270A-5

```

```

Query Match      12.1%; Score 262; DB 3; Length 478;
Best Local Similarity 24.2%; Pred. No. 9.4e-17;
Matches 111; Conservative 86; Mismatches 178; Indels 84; Gaps 18;

```

```

QY 19 IGYIATLTFVGLGMAAPPLDSAVIANGVS-AEVSQVQHLEGMLAKILVREGKV 77
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 60 IAYLIMLFLELAV-LSIVSHVEIVATATGKLAFSDRSKEIKPIENALVKEIVQDQFV 118
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 78 KAGOVLELPTQANAAAGTRNQYVALKAMEAR--LLAE--RDQPSISFAD--LTS 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 119 EKQILLNLTLALGADADQOKTKSSLTLEKRYEITLRAVADRPLIELTKDEFKHA 178
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 131 QRADPMVARAIADQOAFTR-----ROTIQGVDMNAQRLQYSEIEGIDRQT 180
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 179 TEEDKTRIRRLITEGFQFAMQKQYKELALORREAEKQTVLANIRK-----YEGISR-- 230
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 QGLKDLQGLFIEDE-LIDRLKLYDKGLVPRRLALAEAR---AGSISGSIGRLTADRSKA 235
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 231 -----VENERKLXDKLKLFNFSKSTSKHDVLTQENRHEAVNEIAYKSRINEVESDL 281
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 236 VOGASDTOLKVRQIKQEFEEYOSITETRVRLAEYTEKEVNASDQKRKIY--SPVNV 293
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 282 RQAKKEIHLITQLEFRADILKQNV-EAKQSLSELEK-----NEQRLASVIRAFVSG 335
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 294 TAONLREFTEGAVVRAEPLVDIAPEDEAFVIOAHFQPTVDVNMGMVTEVRLPAFHSA 353
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 336 TVQQLKTHVGGVVTAEFLMVAIPEDDVLEVTALQNKDIGFIEVGQDAVIVKVEFPY 392
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 354 GNPDEPRHDVVA-----VADRIADPQKARLFLGIVRVVYKQLPRHLRG--- 398
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 393 -----PYTRYGLMGVKNITTEAIEHPQ-LGLVFNLSIISIDRKT-----LSGKDGK 438
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 399 ---VTAGMPAQVIVPGEPTVLOYLESPLDRLRTMRE 434
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 439 EIELSGMSVTAIEIKGERSVISTLSPLESVSESLRE 477
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 5
US-08-853-659A-38
; Sequence 38, Application US/08853659A
; Patent No. 5925522
; GENERAL INFORMATION:
; APPLICANT: Wong, K.K.; Saifer, J.D.
; TITLE OF INVENTION: A Salmonella Sequence, Methods of Detection
; TITLE OF INVENTION: Of A
; TITLE OF INVENTION: Salmonella Sequence, And Methods of Detection Of Salmonella
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Paul W. Zimmerman

```

```

; ADDRESSEE: Intellectual Property Services
; ADDRESSEE: Battelle Memorial Institute
; ADDRESSEE: PNNL P.O. Box 999
; STREET: Washington Way
; CITY: Richland
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 99352
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.50", 1.44 Mb storage
; COMPUTER: IBM PC/XT/AT
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Processor (WordPerfect 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/853,659A
; FILING DATE: Unknown
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: none
; FILING DATE: n/a
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 425 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-853-659A-38

```

```

Query Match          9.3%; Score 201.5; DB 2; Length 425;
Best Local Similarity 21.0%; Pred. No. 4.6e-11;
Matches 94; Conservative 93; Mismatches 187; Indels 73; Gaps 19;

```

```

QY 7 ORPDNFOAVARIGYGIATLFGVGLGMAFAPLDSAVIANGVSA-EVSODVOHLEGGM 65
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 3 RROSDHLMIT-----IISLTIL-IILTYFIEINSYVHGQGVITTTDNMQLISLKGST 55
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 66 LAKLIVREGKRVAGQVL---FELDPTQANAAAGITRNOYVALKAMEARLLAERDQRPST 122
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 56 IODIYVAEGDTVKKGELLAKVND-----LQKEYQRYRTQGYL--DKDVN-EI 102
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 123 SFPRDLTSQRADPNV--ARAIADQAOFTERRQTIQGOVULMNRQLOXOSETGIDRQT 180
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 103 SFILDKNEGSLTLDGTRSLSNREK--ANIELVHSQIRAKELKTSLSSEISGLQEKL 160
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 181 QGLKQDGFIEDELIDRLKTYDKGLV-----RPRLALEARAGSLSGISGLTADRS 233
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 161 SKRKEALLAEENILISPLTKKGISPYTNFLNKKQAVIKKSIINDIESTI-TLKKDDI 219
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 234 KAVOGASDTQLKVRQIKQEFFEYVSOSITETRV--RLAEVTEKEVVASDAQRIKIVSPV 291
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 220 ELV--VNDIEALNMLRLSLKISKLNQLELVVNSTLKVLEKQINDEED-----IYSPV 271
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 292 NGTQO--NLRFTEGAVVRAERPLVDIAPDEAVIQAHQPTVDVNVHGMVTEVRLPA 349
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 272 DGVLYTKINKSATTHGVYIQADLLFEIKPKVRYMLADVKILPKRYDOIYDEAVKLDVQS 331
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 350 FHSAGNPDPERHDVAADRISDPQKQ-----ARLEFIVRVDVQO-----LPHLRG 397
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 332 I-----IQPKTKSNATINDISPSYEENGTGQIRYKVIADVNVEDDLRWLKP----- 382
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 398 RVTAGMPAQVIVPTGERTVQLYESPL 424
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 383 ----GMTVDASVITGKSHIMEYLLSPL 405
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

```

RESULT 6
US-08-868-699A-2
; Sequence 2, Application US/08868699A
; Patent No. 6204019
; GENERAL INFORMATION:
; APPLICANT: O'Dwyer, Karen
; APPLICANT: Perry, Caroline
; APPLICANT: Warren, Richard L.

```

```

; TITLE OF INVENTION: No. 6204019e1 Compounds
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/868,699A
; FILING DATE: 04-JUN-1997
; CLASSIFICATION: 536
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GM10012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
;
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 796 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-868-699A-2

```

```

Query Match          5.5%; Score 118.5; DB 4; Length 796;
Best Local Similarity 20.6%; Pred. No. 0.0099;
Matches 63; Conservative 52; Mismatches 120; Indels 71; Gaps 11;

```

```

QY 10 TDNFOAVARIGYGIATLFGVGLGMAFAPLDSAVIANGVSAEVSODVOHL----- 61
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 358 TIFQNLFKFESPSGMTATGKLGSEFFDLYSKTI---VVOAFTDKAIRIDEPDKVFR 413
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 62 -----EGMLAKILVREGKRV--AGQVLFELDPTQANAAAGITRNOYV 103
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 414 SVDEKNTAMTHDIVELHETGEPVLLITRTAIAEYSEVLFQMD-----IPNLLI 464
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 104 ALK-AMBARLLAERDQRPSTSFADLTSSORADPNVVARAIADQAOFTERRQTIQGOVULM 162
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 465 AQNVAKKEQMIAEKGQIGSMVTATSMAGRGTDILGEGV-----EALGGVLVII 513
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 163 NAQRLQYQSEIEGIDROTQGIKQDLG-----FIEDELIDRLKTYDKGLVPRPLLA 213
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 514 HE-----HMENSRVDRDLRGSGRGQDPGSSCIYISLDDYLV--KRWSDSNLENNQLYS 566
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 214 LEARAGSLSGISGLTDRSKAVOGASDTQ--TKVRQIKQEFFEVS--GSIETVRVLA 269
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 567 LDAQRLQSQNLNFKRKQVYKQRISEEGVAKREMANEFKSTISIQRLVYEERRVL 626
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 270 EVTEKE 275
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 627 EIDDAE 632
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

```

RESULT 7
US-09-757-014-2
; Sequence 2, Application US/09757014
; Patent No. 6348342
; GENERAL INFORMATION:
; APPLICANT: O'Dwyer, Karen

```


Db 692 INOLEEPLISCAGSADHLLSTVTSISSCIEQLEKSMYSOYLACPEDISGLHSTITLHL 751
QY 330 QPTDVNVHMGWTEVRLPAFHSAGNDPDERHDPVAVADRI SDPQK 376
Db 752 TS---DAIAGATTCLRAP-----PEP-----ADSLTEACKQ 780

RESULT 9

US-09-085-199B-4
; Sequence 4, Application US/09085199B
; Patent No. 6235879
; GENERAL INFORMATION:
; APPLICANT: Hayden, Michael R.
; APPLICANT: Hackam, Adigail
; APPLICANT: Hug, A.H.M. Mahbulul
; APPLICANT: Chopra, Vikramjit Singh
; APPLICANT: Kalchman, Michael
; TITLE OF INVENTION: Apoptosis Modulators That Interact with the
; TITLE OF INVENTION: Huntington's Disease Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Opedahl & Larson
; STREET: PO Box 5270
; CITY: Ftisco
; STATE: CO
; COUNTRY: USA
; ZIP: 80443-5270
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Kb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS DOS 5.0
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/085,199B
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Larson, Marina T.
; REGISTRATION NUMBER: 32038
; REFERENCE/DOCKET NUMBER: UBC.P-013052
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (970) 668-2050
; TELEFAX: (970) 668-2052
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 914
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: no
; ORIGINAL SOURCE:
; ORGANISM: human
; FEATURE: Huntington-interacting protein
US-09-085-199B-4

Query Match 5.2%; Score 112; DB 4; Length 914;
Best Local Similarity 20.1%; Pred. No. 0.051;
Matches 82; Conservative 72; Mismatches 131; Indels 122; Gaps 20;

QY 47 NGVVAEVSQDVQHLE---GMLAKILVREGEKVKAGCVLEFELDPQTQANAAGITRNOYV 103
Db 243 NGVNKDEKHILIERLYREISGLKAOI---ENMKTESQRVYLQIKGHVSELADLAEOQHL 299
QY 104 ALKAME--ARLLAERDQPSISFPADLTSGRADPVARAIAD---QKQTFERR----- 152
Db 300 ROQADDCERFAELD-----ELKROREDTEKARSLSEIERKQANEQRYSKLKE 350
QY 153 ---STIGOVILM--NAO-----RLQYSEIIGIDRQTOGLKDOIGFIED-----ELI 195
Db 351 KYSELVQNHADLLKRNNAEVNTQVSMARQAYD--LEREKKELEDSEIERISDQGRKTQSOQL 409
QY 196 DLRLYDKGLVPRRLALLEARAGSLSGS-----IGRLADRASKAVQASDTOLK 245

Db 410 EYLESLKQELGTSQR--ELQVLOGSLETSTAQSEANMAAEFAELEKERDLSVSGAHRNEE 467
QY 246 VROIKQEFFEFGVOSITETFEVRLAEVTEKEVVASDAQKRIKIVSPVNGTAONIRFFEGA 305
Db 468 LSAIRKE-----LQDTQIKLASTEESMCQAKDKRMILY---GSKRAAEQVLIQDA 515
QY 306 VRAAE-PLVDA-----PED-----EAEVIOAHF 329
Db 516 INOLEEPLISCAGSADHLLSTVTSISSCIEQLEKSMYSOYLACPEDISGLHSTITLHL 575
QY 330 QPTDVNVHMGWTEVRLPAFHSAGNDPDERHDPVAVADRI SDPQK 376
Db 576 TS---DAIAGATTCLRAP-----PEP-----ADSLTEACKQ 604

RESULT 10

US-08-642-846-2
; Sequence 2, Application US/08642846
; Patent No. 5886151
; GENERAL INFORMATION:
; APPLICANT: HOSTETTER, MARGARET K.
; APPLICANT: GALE, CHERYL A.
; APPLICANT: BENDEL, CATHERINE M.
; APPLICANT: TAO, NIAN-JUN
; APPLICANT: KENDRICK, KATHLEEN
; TITLE OF INVENTION: CANDIDA ALBICANS GENE, INTEGRIN-LIKE
; TITLE OF INVENTION: PROTEIN, ANTIBODIES, AND METHODS OF USE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MUEITING, RAASCH, GEBHARDT & SCHWAPPACH, P.A.
; STREET: 119 NORTH FOURTH STREET, SUITE 203
; CITY: MINNEAPOLIS
; STATE: MINNESOTA
; COUNTRY: USA
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/642,846
; FILING DATE: 03-MAY-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MUEITING, ANN M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 110.00280101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-305-1217
; TELEFAX: 612-305-1228
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1664 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-642-846-2

Query Match 5.0%; Score 109; DB 2; Length 1664;
Best Local Similarity 19.5%; Pred. No. 0.24;
Matches 80; Conservative 61; Mismatches 124; Indels 146; Gaps 20;

QY 92 NAAAGITRNOYVALKAMEARLLAERDQPSISFPADLTSGRADPVARAIADDEQAQTFER 151
Db 974 NAKKGVGTQDEYI---NAKLVQDKPKKNSI-----VTDPEDRYEEL 1010
QY 152 ROT-----IQGVLDLMAHQRLQYSEIIGIDRQTOGLKDOIGFIEDLIDLRK- 199
Db 1011 QQTASINHATIDSSISYGRPDSISTDMLEPLISD--ELKKRPRTALLSADRLFMQEVNHPILRSN 1069

QY 200 -----LYDKGLVPRRLLEAARAGSLSSIGRLTADRSKAVOGASTOLKYROI 249
Db 1070 SLYVHPGAGATNSMLPEPDFELINSPARVNSNDNV-----AISGNAST-ISENQL 1122
QY 250 KQEFEEQ-VQSQITE-----TRVLAETVEKEVASDAQ 282
Db 1123 DMNFDQATIGQIOEQPASKSANTVGRGDDGLASABETPTPKKSISSKPAKLSSAS 1182
QY 283 KR--IKIVSPVNGTAONLRFETEGAVVRAEPLVDIAPEDEAVIQAHQPTVDN--V 337
Db 1183 PRKSPIKIGSPV-----RVKKNGSIAGIEPIPKATHKP-----KKSFGNGEISNHKV 1230
QY 338 HMGVTE-----VRLPFIHSAGN--PDP-----ERHD-- 362
Db 1231 RDGGISPSGSEHQHNPMSVSPSOYTDATSTVPDENKDVQHKPREKOKKHHNRHHN 1290
QY 363 -----PAVADRISDP--QKQARLF--LGIVRVYKOLPRHLGRVT 400
Db 1291 HHKQKTDIPGVVDEIDPDVGLQERGLFFRVYLGIKINLPDINTH-KGRFT 1340

RESULT 11
US-09-264-604-2
; Sequence 2, Application US/09264604
; Patent No. 6346411
; GENERAL INFORMATION:
; APPLICANT: HOSTETTER, MARGARET K.
; APPLICANT: GALE, CHERYL A.
; APPLICANT: BENDEL, CATHERINE M.
; APPLICANT: TAO, MIAN-JUN
; APPLICANT: KENDRICK, KATHLEEN
; TITLE OF INVENTION: CANDIDA ALBICANS GENE, INTEGRIN-LIKE
; TITLE OF INVENTION: PROTEIN, ANTIBODIES, AND METHODS OF USE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MUEITING, RAASCH, GEBHARDT & SCHWAPPACH, P.A.
; STREET: 119 NORTH FOURTH STREET, SUITE 203
; CITY: MINNEAPOLIS
; STATE: MINNESOTA
; COUNTRY: USA
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/264,604
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/642,846
; FILING DATE: 03-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: MUEITING, ANN M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 110,00280101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-305-1217
; TELEFAX: 612-305-1228
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1664 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-264-604-2

Query Match 5.08; Score 109; DB 4; Length 1664;
Best Local Similarity 19.58; Pred. No. 0.24;

Matches 80; Conservative 61; Mismatches 124; Indels 146; Gaps 20;

QY 92 NAAGTIRNOGVALKANEARLLARDQRPSTSPFADLTSGQADMMVRAITDEQAQFTER 151
Db 974 NAKKGVTODEYT-----NAKLVDQKPKKNSI-----VTPDREYEL 1010
QY 152 ROT-----IOGVDLMNAQRLQYOSEIGIDRQTOGLKDQGLFIDELEIDRK- 199
Db 1011 QQTASINHATIDSSITYGRPDSISTDMLPYLSD-ELKKRPITLLSADRLFMQGEVHPLRSN 1069
QY 200 -----LYDKGLVPRRLLEAARAGSLSSIGRLTADRSKAVOGASTOLKYROI 249
Db 1070 SLYVHPGAGATNSMLPEPDFELINSPARVNSNDNV-----AISGNAST-ISENQL 1122
QY 250 KQEFEEQ-VQSQITE-----TRVLAETVEKEVASDAQ 282
Db 1123 DMNFDQATIGQIOEQPASKSANTVGRGDDGLASABETPTPKKSISSKPAKLSSAS 1182
QY 283 KR--IKIVSPVNGTAONLRFETEGAVVRAEPLVDIAPEDEAVIQAHQPTVDN--V 337
Db 1183 PRKSPIKIGSPV-----RVKKNGSIAGIEPIPKATHKP-----KKSFGNGEISNHKV 1230
QY 338 HMGVTE-----VRLPFIHSAGN--PDP-----ERHD-- 362
Db 1231 RDGGISPSGSEHQHNPMSVSPSOYTDATSTVPDENKDVQHKPREKOKKHHNRHHN 1290
QY 363 -----PAVADRISDP--QKQARLF--LGIVRVYKOLPRHLGRVT 400
Db 1291 HHKQKTDIPGVVDEIDPDVGLQERGLFFRVYLGIKINLPDINTH-KGRFT 1340

RESULT 12
US-08-816-693A-51
; Sequence 51, Application US/08816693A
; Patent No. 5874241
; GENERAL INFORMATION:
; APPLICANT: Takahashi, Joseph S
; APPLICANT: Turek, Fred W
; APPLICANT: Pinto, Lawrence H
; TITLE OF INVENTION: Clock Gene and Gene Product
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Rocky, Milnamow & Katz
; STREET: Two Prudential Plaza, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/816,693A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5874241Inrup, Thomas E
; REGISTRATION NUMBER: 33,268
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5400
; TELEFAX: 312-616-5460
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 747 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-816-693A-51

Search completed: July 19, 2002, 10:41:39
Job time: 39 sec

```

RESULT 15
US-08-179-738-5
; Sequence 5, Application US/08179738
; Patent No. 5578462
; GENERAL INFORMATION:
; APPLICANT: Seizinger, Bernd R.
; APPLICANT: Kley, Nikolai A.
; APPLICANT: Bianchi, Albert B.
; TITLE OF INVENTION: No. 5578462el NF2 Isoforms
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reed & Robins
; STREET: 635 Bryant Street
; CITY: Palo Alto
; STATE: California
; COUNTRY: U.S.A
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/179,738
; FILING DATE: 10-JAN-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Robins, Roberta L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 5998-0017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 591 amino acids
; TYPE: amino acid
; TOPOLOGY: 1linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: murine
; US-08-179-738-5

```

Query Match 4.9%; Score 106.5; DB 1; Length 591;
Best Local Similarity 21.6%; Pred. No. 0.087;
Matches 69; Conservative 54; Mismatches 117; Indels 79; Gaps 14;

```

QY 124 FPAADLSQARADPVVARAIADDEQAQFTERRQTIGQYDLNMAQ---RLQYQSEIEGIDRQ 179
DB 283 FKFDSSKLRNKKLILQLCIGNHDLFMRRKRAADSEVQOMKAAQAREKARQMERQRLARE 342
QY 180 TQGLKDLQGFIEDEL-IDLRKLYDKGLVPPRRLALEARAGSLSGSIGRTADRSKAV-Q 237
DB 343 KQ-MREEAERTRELRLRLQMKKEATMANEALMRSEETADLAER-AQTTEEAQILAQ 400
QY 238 GASDTQKVAQIQKEFFEOVSOSITETRVRLAE--VTEKEV---VASDAQKRIKIVSPV 291
DB 401 KAAEAQCEQMORIKATAIR-----TEEEKRLMQKYLEAEVLALNMAEESERRAKEADQL 454
QY 292 NGTAQNLRFTEGAVVRAAEPLVIDA-----PEDEAFVIGA-----HFQ 330
DB 455 KQDLQEAR---EARRAKOKLEIAIKTPYPMPNPPIPPPLPIPSFDIIADSLSDFEK 510
QY 331 PTDVNVVHM-----GAVTEVR-----LPAFHSAGNDPPE 359
DB 511 DTDMKRLSMIEIEKXEYMEKSKHLOQNLNETLIEALKLKREFTALDVLHSESS---D 567
QY 360 RHPDPAVADRISDPQOKAR 378
DB 568 RGGPSSSKHDTIKKPKQAQGR 586

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2002, 10:42:14 ; Search time 32.83 Seconds
(without alignments)
1273.191 Million cell updates/sec

Title: US-09-913-414-5

Perfect score: 1 MKPKIQRPDNEQAVARIG.....VLQYLFSPLRDRLRTTMRRE 435

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|--------|--------------------|
| 1 | 1984.5 | 91.7 | 436 | 2 | E87374 | RsaA secretion sys |
| 2 | 589 | 27.2 | 474 | 2 | F98310 | rhizobium secret |
| 3 | 584 | 27.0 | 437 | 2 | AE2972 | secretion protein, |
| 4 | 583 | 26.9 | 439 | 2 | C96003 | protein secretion |
| 5 | 568 | 26.2 | 434 | 2 | AB3099 | hypothetical prote |
| 6 | 568 | 26.2 | 436 | 2 | G98187 | rhizobium secret |
| 7 | 564 | 26.1 | 443 | 2 | B83231 | metalloprotease |
| 8 | 543.5 | 25.1 | 473 | 2 | C95953 | hypothetical prote |
| 9 | 542.5 | 25.1 | 437 | 2 | H97683 | pse protein (Y127 |
| 10 | 542.5 | 25.1 | 437 | 2 | A12908 | HLVD family secret |
| 11 | 530.5 | 24.5 | 511 | 2 | D97753 | alkaline protease |
| 12 | 519.5 | 24.0 | 511 | 2 | D71687 | alkaline protease |
| 13 | 514.5 | 23.8 | 432 | 2 | S26697 | apre protein - pse |
| 14 | 514.5 | 23.8 | 432 | 2 | G83489 | alkaline proteinas |
| 15 | 506.5 | 23.4 | 448 | 2 | S12526 | metalloprotease |
| 16 | 476.5 | 22.0 | 442 | 2 | AE0477 | HLVD family secret |
| 17 | 471.5 | 21.8 | 437 | 2 | BA4933 | hasa export system |
| 18 | 438.5 | 20.3 | 452 | 2 | AB2965 | secretion protein, |
| 19 | 438.5 | 20.3 | 497 | 2 | C98318 | alkaline proteinas |
| 20 | 368.5 | 17.0 | 387 | 2 | AG0835 | probable type I se |
| 21 | 368 | 17.0 | 481 | 2 | D82381 | secretion protein, |
| 22 | 367 | 17.0 | 395 | 2 | D83412 | probable secretion |
| 23 | 329 | 15.2 | 475 | 2 | B81047 | secretion protein, |
| 24 | 325 | 15.0 | 391 | 2 | D85547 | probable membrane |
| 25 | 325 | 15.0 | 391 | 2 | H90696 | probable membrane |
| 26 | 314 | 14.5 | 473 | 2 | F82561 | hemolysin secretio |
| 27 | 308.5 | 14.2 | 477 | 2 | S48045 | toxin aprxII secre |
| 28 | 288.5 | 13.3 | 478 | 1 | LEECDC | hemolysin export s |
| 29 | 288 | 13.3 | 507 | 2 | AD2140 | hypothetical prote |

| | | | | | | |
|----|-------|------|-----|---|--------|---------------------|
| 30 | 267.5 | 12.4 | 477 | 2 | B61378 | leukotoxin secreti |
| 31 | 267.5 | 12.4 | 478 | 2 | S10058 | hemolysin secretio |
| 32 | 266 | 12.3 | 440 | 1 | BVBRCD | cyad protein - Bor |
| 33 | 264 | 12.2 | 438 | 2 | T03513 | probable secretion |
| 34 | 262 | 12.1 | 478 | 2 | E43599 | toxin aprxII secret |
| 35 | 261.5 | 12.1 | 454 | 2 | G81828 | probable periplasm |
| 36 | 260 | 12.0 | 500 | 2 | AB2517 | hemolysin secretio |
| 37 | 251.5 | 11.6 | 479 | 2 | T00229 | hemolysin secretio |
| 38 | 251 | 11.6 | 467 | 2 | G82198 | rxn toxin transpor |
| 39 | 248 | 11.5 | 498 | 2 | AH2468 | hypothetical prote |
| 40 | 234 | 10.8 | 478 | 2 | D30169 | leukotoxin secreti |
| 41 | 215.5 | 10.0 | 583 | 2 | S75805 | hemolysin secretio |
| 42 | 203.5 | 9.4 | 425 | 2 | AB1018 | probable type-I se |
| 43 | 201.5 | 9.3 | 425 | 2 | T14873 | HLVD secretion pro |
| 44 | 196 | 9.1 | 455 | 2 | D95923 | probable OMA famil |
| 45 | 194 | 9.0 | 512 | 2 | AB2047 | hypothetical prote |

ALIGNMENTS

| Query Match | Best Local Similarity | 91.7% | Score 1984.5; | DB 2; | Length 436; |
|-------------|-----------------------|---|---------------|------------|-------------|
| Matches | 408; | Conservative | 3; | Mismatches | 24; |
| | | | | Indels | 1; |
| | | | | Gaps | 1; |
| QY | 1 | MKPKIQRPDNEQAVARIGYIALTFVGLIGNAFAPLDSAVIANGVSAEVS-QDVQ | 59 | | |
| Db | 1 | MKPKIQRPDNEQAVARIGYIALTFVGLIGNAFAPLDSAVIANGVSAEGRKTVQ | 60 | | |
| QY | 60 | HLEGMGLAKIIVREGKAKQGVLELDPQANNAAGITRQYVALKMEKRLAERDQR | 119 | | |
| Db | 61 | HLEGMGLAKIIVREGKAKQGVLELDPQANNAAGITRQYVALKMEKRLAERDQR | 120 | | |
| QY | 120 | PSISPADLTQSORDPWARAIADBOAQFERRROTIGOGVLMNAORQYQSEIIGIDRQ | 179 | | |
| Db | 121 | PSISPADLTQSORDPWARAIADBOAQFERRROTIGOGVLMNAORQYQSEIIGIDRQ | 180 | | |
| QY | 180 | TQGLKQDQGFTEDELIDRLKLYDKGLVPRRLALLEARAGSLSGISGRITADRSKAVQGA | 239 | | |
| Db | 181 | TQGLKQDQGFTEDELIDRLKLYDKGLVPRRLALLEARAGSLSGISGRITADRSKAVQGA | 240 | | |
| QY | 240 | SDTQLKQVQIOEFPEVQVQSITETRVRLAEVTEKEVVASDAQRIKIVSVNGTAQNL | 299 | | |
| Db | 241 | SDTQLKQVQIOEFPEVQVQSITETRVRLAEVTEKEVVASDAQRIKIVSVNGTAQNL | 300 | | |
| QY | 300 | FETGCAVVRRAEPLVDIAPDEDAFVIOAHFOPTDVNVHGMVTEVRLPAPHSGNPPE | 359 | | |
| Db | 301 | FETGCAVVRRAEPLVDIAPDEDAFVIOAHFOPTDVNVHGMVTEVRLPAPHSGNPPE | 360 | | |
| QY | 360 | RHDPVAVADRISDPQKARLFLGIVRVQKQPLPHLRGRVATAGPAQVIVTGETVLYQ | 419 | | |
| Db | 361 | GTIQSLSDRISDPQKARLFLGIVRVQKQPLPHLRGRVATAGPAQVIVTGETVLYQ | 420 | | |

```
Oy      420 LFSPDLRTMTREE 435  
          |||||  
Db       421 LFSPDLRTMTREE 436  
  
RESULT   2  
F98310  
rhizobioin secretion protein rspe (AF141932) [imported] - Agrobacterium tumefaciens (str.  
C.Species: Agrobacterium tumefaciens  
C.Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002  
A.Accession: F98310  
R.Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
A.; Liu, F.; Woliam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Marxeltz, B.;  
Science 294, 2323-2328, 2001  
A.Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
A.Reference number: A97359; PMID:11743194  
A.Accession: F98310  
A.Status: preliminary  
A.Molecule type: DNA  
A.Residues: 1-474 <XUR>  
A.Cross-references: GB:AE007870; PIDN:AAK90008.1; PID:g1515979; GSPDB:GNO0170  
C.Genetics:  
A:Gene: AGR_L_2881  
A:Map position: linear chromosome
```

| | | | | |
|-----------------------|------------------|--------------------|-----------|-------------|
| Query Match | 27.2%; | Score 589; | DB 2; | Length 474; |
| Best Local Similarity | 31.9%; | Pred. No. 3.1e-29; | | |
| Matches 140; | Conservative 85; | Mismatches 210; | Indels 4; | Gaps 2; |

[illegible]

RESULT 3
AE2972
secretion protein, HlyD family [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AE2972
R:Wood, D.W.; Seubdal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.;
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kulyavin, T.; Levy, R.; Li, M.; McCellan,
; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AE2972

A:Cross-References: GB:AE008669; PIDN:AAU4195.1; PID:q17741773; GSPDB:GN00187
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: pISE
A:Map position: linear chromosome

| | | | | |
|-----------------------|--------------|-------------------|------------|-----|
| Query Match | 584 | DB 2 | Length | 437 |
| Best Local Similarity | 31.98 | Pred. No. 5.7e-29 | | |
| Matches 139 | Conservative | 84 | Mismatches | 209 |
| | | | Indels | 4 |
| | | | Gaps | 2 |

[illegible]

RESULT 4
C96003
protein secretion protein, *HiYD* family [imported] - *Shimorhizobium meliloti* (strain 10
C.Species: *Shimorhizobium meliloti*
C.date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C.accession: C96003
R.Finan, T.M., Weidner, S., Wong, K., Buhmester, J., Chain, P., Vorholter, F.J., Herr
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A.Title: The complete sequence of the 1,663-kb pSymb megaplasmid from the N2-fixing e
A.Reference number: A95842; MUID:21396508; PMID:11461431
A.Accession: C96003
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-439 <KUR>
A.Cross-references: GB:A591985; PIDN:CAC49691.1; PID:g15141178; GSPDB:GN00167
A.Experimental source: strain 1021, megaplasmid pSymb
R.Gallbert, F., Finan, T.M., Long, S.R., Puhler, A., Abola, P., Ampe, F., Barloy-Hubba
pella, D., Chain, P., Cowie, A., Davis, R.W., Dreano, S., Federspiel, N.A., Fisher, R.
L., Hyman, R.W., Jones, T.
Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalmun, S.; Keeling, D.H.; Kiss, E.; Komp, C.; Lelaure, A.; Neubalt, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: prsB; SMB21465
A:Genome: plasmid

| | | | | |
|---------------------------|-------|--------------------|-----------|-------------|
| Query Match | 26.9% | Score 583; | DB 2; | Length 439; |
| Best Local Similarity | 31.9% | Pred. No. 6.6e-29; | | |
| Matches 134: Conservative | 84: | Mismatches 198: | Indels 4: | Gaps 4 |

| | | | |
|----|-----|--|-----|
| Oy | 19 | IGYVIALTEFV-GLLGNAPAFPLDSAVTANGVSXNVS-DOVHLEGGMLAKITVRECK | 76 |
| Dd | 21 | IGVSVLALALVAGVGGMATTELSSAIYGVGVYVDVVKKVQHLTGIVGELLKKEEDR | 80 |
| Oy | 77 | VKAGOVLFELDPTOANAAGITTRNOYVALKAMEARLLAERPORSISHPALUTSORADPM | 136 |
| Dd | 81 | VEAGOVILRIDGTYYRANLALITESTIAQFYARRRRRLQAEIRRGASFELEDLAEFIGTA | 140 |
| Oy | 137 | VARAIADOAOFTERRROTIQOQVDMNNAQLQYOSEIIEGDROTQGLKDOLGFIETDELID | 196 |
| Dd | 141 | AAKLIEGEORLFAFSRSLSGMSKGOLDLSRKAQDLADEVEGLTVQJNATAIEEALKLIAEELTG | 200 |
| Oy | 197 | LRLKYDGLVPRPRLMLEARAGSISGSIGRLTDRSKNAVGAADTLQKVNQIOEFFEQ | 256 |
| Dd | 201 | VDLSLEGOGVLVMORTYLTKKRORALEEGGRHHIARAQOKSSSEIIDQTILLDEDRSE | 260 |
| Oy | 257 | VOSOTTERVRVALEVEKEVASDAQRIKIVSPVNGYAQMRFETBEGAVVRAEPVLVDI | 316 |
| Dd | 261 | ISKELTEDVEAKIAIEEBERTATDQLRDLITARLSGIYLOLAHTYNGVINPEBTIMLV | 320 |
| Oy | 317 | APDEBAFYQAHFOPDVNVHMGMATEVURLPAFHSAQNPPERRHDVYAVADRISDPQK | 376 |
| Dd | 321 | VPEADEDLVEAKVATHDIDQIRVGOSEVIERPSAFNQRTTPEVEAEVVTVAPDLVTDBERG | 380 |
| Oy | 377 | ARLFLGIVRVNVKQLPRILRG-RVYTAGMPAYUIPYTGERTYLOLFLFSRLNTLETTRYEE | 435 |
| Dd | 381 | ASYPLRLRPRAESI-AKLKLSIYIPGPAPAEVFIKIADRYISLYTPLDLDMNHAREP | 439 |

```

Query Match      26.2%; Score 568; DB 2; Length 434;
Best Local Similarity 31.5%; Pred. No. 5,6e-28;
Matches 138; Conservative 79; Mismatches 201; Indels 20; Gaps 5;
07 11 DNFAQVARI-----GYGIIATLFFVGLGCAAPPLDSAVIANGCVSAEVS-QDVQHL 61

```

| | | | |
|----|-----|---|-----|
| Dh | 2 | DEMÖLKRISIRSHLLVGLGFLTLTGv- FCGHAvSTElVGAVIAGSLVETSLKKVdHP | 60 |
| Qy | 62 | EGGMLKTLVREGEKVKAGOVLFELDPTQANAAGITRNOVVALKAMEARLLAERDQPS | 121 |
| Dd | 61 | VGVVvSELMvNRGdGVKvKAGDVvMRIDATvMTRANLAIvVKSIdQfTARARALEGEDRAAS | 120 |
| Qy | 122 | ISFPdLTSQrADpVvARAIADeQAQfTERRQTLQGOvDLvMNAQRLOyQSEIEGIDROTQ | 181 |
| Dd | 121 | VYFPOSLDRAGDAEvLLMMNAEQRLYEDRAVRvRSKKROLEQvRQRLREISGLEAERA | 180 |
| Qy | 182 | GLKDOLGFIEdELDLRLKLYOKGLvPRPRLALeARASLSGSIGRLADSKvAGQASD | 241 |
| Dd | 181 | ANVRÖGMvDESLIFRSIQERGSLDKSRSLTEKQADVIDODIGRLAGIAGIEAKTISE | 240 |
| Qy | 242 | TÖLKRAQIKOEFEEOVSOSITETRVRLAEvTEKEVvASDAQRIKvISPVNGTQANLRFf | 301 |
| Dd | 241 | TALÖTLQIDEOMSEEVGSLSLRMDARIGELYERRAAEDQLKRVIDLARPOGVvHQLAVH | 300 |
| Qy | 302 | TEGAVvRAAEPLvDIAPEDeAEvIQAHFQPTDvDvNHvMGvTEvRLPAFvHSGvNDPvERH | 361 |
| Dd | 301 | TVGAVIARGEQIMvIPEvDKLVLvEAKvAPQIDQIYFGvTNLvFFSvNQKTPv----- | 355 |
| Qy | 362 | DEvAVvADvRIS-----DPQÖARLEFLGIVRVdYKQLPvHLRGvTAgMAEQvIVPvGGETv | 416 |
| Dd | 356 | ELTGTvERISADvTVDOFGASyYvLVRvATvSOEQIKRIGESvLPMGMvEAvEAFITvGERSv | 415 |
| Qy | 417 | LÖYLFSPvLBDTLRTvMRE | 434 |
| Dd | 416 | LSYFLKPLDvDQANvRTFQ | 433 |

[illegible]

242 TOLKVRKQKOEFFEOVSQSTTETRVRLAEVTEKEVVASDAQKRIKIVSPVNGTAQNLRF 301
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
243 TALQIILQIDPQMSSEVGSDLREMDARIGEVERRAAEDQKRDILAPADGVHQLAVH 302
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 302 TECAVRAAEPLVDIAPEDEAFVIOAHFOPTDVNDVNHMGVTEVRLPAFHSAGNPPPERH 361
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 303 TVGVGIAPGQIMMIVIEVDKRLVEAKVAQODIDQIYFGVNTLRSFAFQKQTP----- 357
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 362 DPVAVADRIS-----DPQKARLEFLGIVRVDKOLPPLHGRVTAQMPAQVIVPTGERTV 416
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 358 EITGYVERISADTVQDRIGASVYLVKVAISQDIKRLGEFLMPCGMFPAFAFITGERSV 417
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 417 LQYLFSPRLDRLTMR 434
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 418 LSYFLKPLDQANTRFRQ 435
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

RESULT 7
B83221
metalloproteinase secretion protein PA3405 [imported] - Pseudomonas aeruginosa (strain H
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: B83221
R:Stover, C.K.; Pham, X.Q.; Ewlin, A.L.; Muzoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lartig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; M0ID:20437337
A:Accession: B83221
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-443 <STO>
A:Cross-references: GB:AE004761; GB:AE004091; NID:9949533; PIDN:AAG06793.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: hase; PA3405

Query Match 26.1%; Score 564; DB 2; Length 443;
Best Local Similarity 32.5%; Pred. No. 1e-27;
Matches 18; Conservative 87; Mismatches 189; Indels 10; Gaps 4;
QY 17 ARIGYIATLFVGLGMAFAPLDSAVIANGVSAEVSOD-VQHLBGMLAKILVEGE 75
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 25 SRLGWIIVLGFVGLMAGLPLDKGVSGTVWAGSKRAVGHPTGGLRVHRIHVEGE 84
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 76 KVAAGVLELPTQANAAAGITRNQYVALKAMEARLLAERDQPSISFPADLTQSRADP 135
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 85 RVEAGVILEMDATQARAQADGLFAQYLAALASLARLSAERDEKARLEFPALIELA-LDDP 143
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 136 MVARAIDEDQAFTERROTTOGOVDLMNAQRLQYQSEIEGIDRTOGLKQDLGFIEDLI 195
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 144 RLPTLLDQOQKLDHSRRRLALELDGLAETVYVAGSOAQDLDQAALRSKEQRAALEDR 203
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 196 DLRLYDKGLVPRRLALAEARA-----GSLGSGITRLTADRSKAVOGASDTOLVKROIQ 251
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 204 GLRQLASGEVYPRNRLDSERLLAQVNGELIAGDLSGSTRLT-----ELRLMAQRR 259
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 252 EFFEQVQSOSTETETRVRLAEVTEKEVVASDAQKRIKIVSPVNGTAQNLRFTEGAVRRAE 311
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 260 KFOREVBASLADQAVRAEELNRLASARFDLANSEVRAPVAGLVGQEVTEGEGVIAFGQ 319
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 312 PLVDIAPDEAFVIOAHFOPTDVNDVNHMGVTEVRLPAFHSAGNPPPERHPPVAVADRIS 371
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 320 QLMELPRLDRLTMR 434
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 372 DPQKARLEFLGIVRVDKOLPPLHGRVTAQMPAQVIVPTGERTVLOLFSPLDRLTMR 431
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 380 DERSEAPYIRIVRINVEGEGVRLAGLEIRPGMPEAFVRSERSLANLTFKPLADRLTMR 439
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 432 MRE 435
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

Db 440 LGEE 443
RESULT 8
C95953
hypothetical protein expD2 [imported] - Sinorhizobium meliloti (strain 1021) megaplas
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: C95953
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Her
proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing e
A:Reference number: A95842; M0ID:2136508; PMID:11481431
A:Accession: C95953
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-473 <KUR>
A:Cross-references: GB:AL51985; PIDN:CAC49291.1; PID:91514077; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid pSymB
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hu
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; LeJau
hebaul, P.; Vandendol, M.; Vorholter, F.J.; Weidner, S.; Weiss, D.H.; Wong, K.; Yeh,
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; M0ID:21368234; PMID:11474104
C:Contents: annotation
C:Genetics:
A:Gene: expD2; SMB21315
A:Genome: plasmid

Query Match 25.1%; Score 543.5; DB 2; Length 473;
Best Local Similarity 28.8%; Pred. No. 2.1e-26;
Matches 126; Conservative 97; Mismatches 207; Indels 7; Gaps 3;
QY 4 PKIQRPDNEQAVAR---IGYIATLFVGLGMAFAPLDSAVIANGVSAEVS-QDV 58
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 39 PEPDLSNDNTHSPIRLVINGLTIIVAFGFGFMARSTELSSASVSGTIVDSKRTV 98
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 59 QHLEGMLAKILVEGEKVAAGVLELPTQANAAAGITRNQYVALKAMEARLLAERDQ 118
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 99 SHEFGVGLRILVQEGHVAPEGPLMKLEDTARSDLAQASRVGLIAKLARLARRAG 158
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 119 RPSISFPADLTQSRADPVAARAIDEDQAFTERROTTOGOVDLMNAQRLQYQSEIEGIDR 178
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 159 LHAVSFPADLVSE--GERAADAVTAEKAFPEKRSKAESESLAIQRTIETSEKAKSLTA 216
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 179 QTOGLKQDLGFIEDLILRLYDKGLVPRRLALAEARAGSISGITRLTADRSKAVOG 238
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 217 QLAQTDQIQLMNBQRAIATLVKAPAFQSKLIEIDARLSLAEATRGELAGDAQAEKA 276
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 239 ASDTQLKVRQKOEFFEOVSQSTTETRVRLAEVTEKEVVASDAQKRIKIVSPVNGTAQNL 298
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 277 MAGAELLLIGIESFOSEIAGEITTALEAEVBQRTAAEDVLRLEITRAPQGIYANI 336
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 299 RFEFGAVVRAAEPLVDIAPDEAFVIOAHFOPTDVNDVNHMGVTEVRLPAFHSAGNPP 358
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 337 QLRTPGSAVTPGQPLDIPDEDEPLVBMHVSSTDIDISITIGSSQIRLTAVNRSNHLPL 396
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 359 ERHDPVAVADRISDPQKARLEFLGIVRVDKOLPPLHGRVTAQMPAQVIVPTGERTVLO 418
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 397 EKGATYIAAAGSDVDEKSVAVAFVAREVTPESLANPDIRLPCMPAEVLIVHKSFAID 456
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY 419 YLFSPRLDRLTMR 435
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 457 YLVAPVSDSEFNARRED 473
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
RESULT 9
H97683
prfB protein (Y12758) [imported] - Agrobacterium tumefaciens (strain C58, Cereon)

C.Species: Agrobacterium tumefaciens
C.Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C.Accession: H97683
R.Goodner, B.; Hinkle, G.; Gatting, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Marxelz, B.
Science 294, 2323-2328, 2001
A.Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A.Reference number: A97359; PMID:11743194

A.Status: preliminary
A.Accession: H97683
A.Molecule type: DNA
A.Residues: 1-437 <KUR>
A.Cross-references: GB:AEO07869; PIDN:AAK8425.1; PID:g15157919; GSPDB:GN00169
A.Genetic:
A:Gene: AGRC_4904
A.Map position: circular chromosome

| | Query Match | 25.1% | Score 542.5; | DB 2: | Length 437; |
|----|---|--------|--------------------|-----------------|------------------------------------|
| | Best Local Similarity | 30.8%; | Pred. No. 2.1e-26; | Mismatches 130; | Conservative 83; Indels 5; Gaps 2; |
| OY | 15 AAARIGVGIATLFVGILGMAAFAPLDSAVIANGVSAE-VSODVOHLEGMLAKTVRE | 73 | | | |
| Dd | 20 AYVVVLDELGL---VGGGMAAFPAKLGAAYATGRVVEGNSKKIQLHSISSEINAVE | 75 | | | |
| OY | 74 GERVKAGQVFLELDPTQANNAAGITRNQYVALKAMEARLLAEERDOPRSIPFADLTSSRA | 133 | | | |
| Dd | 76 GDPRVAAGQILLRLSATYYVANLSIENTLAOLYSRRALRAFIAEPSTTYEDTLALS | 135 | | | |
| OY | 134 DPWVARAIADAQAFTERROTIQGOVDLMNAORLOYOSEIEGDIDROTGLKDOLFIDE | 193 | | | |
| Dd | 136 SKSAKFIFDESQLMFNRNALIKMKROLATRKIQLADEARGDLDOVEATEENELATVED | 195 | | | |
| OY | 194 LIDRLKYDKGYLPFRPLLALKEASGSISGRFLPADRSKAAYCASPOLKVRYRKOF | 253 | | | |
| Dd | 196 VSKTDDLLKGKVTLYLOLNLRSLKROLSNEGOGGTARAQCIVGKISELDILOLDED | 255 | | | |
| OY | 254 FEVSOSITEFTVRVLAETFEKEEVVASDAOKRIKTIVSPNGTAONTRFFEGAVBRAEPL | 313 | | | |
| Dd | 256 KSEVTMDITSIEANTVAEEERLATRDOLDIRSPIAGRITYOHSVHNINGVIQGEYL | 315 | | | |
| OY | 314 VDIAPEDEAFEIOAHFOPTDVNVHMGMVTEURLPAPHSGNPDERHDPPVAADRISP | 373 | | | |
| Dd | 316 MLVVPKDDLAIENAITRPDIQIYGOPYVRFETAFNQSFPDLSAEYAAPPDLOTDS | 375 | | | |
| OY | 374 QKOARELGIVARDVKOLPHLKRGRTAVGMPAQYIVPGERVLYOLSPLDTLTTHR | 433 | | | |
| Dd | 376 RICTSYVYLIIRNKMGMLPGSKLYFGMFAEFVFIQTSERSVLSFYFKPFODRLKTFV | 435 | | | |
| OY | 434 EE 435 | | | | |
| Dd | 436 QE 437 | | | | |

RESULT 10
Al2908
HYD family secretion protein [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C.Species: Agrobacterium tumefaciens
C.Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C.Accession: AI2908
R.Wood, D.M.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillett, W.J.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCell
. ; Karpi, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A.Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A>Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A.Reference number: AB2577; PMID:11743193
A.Accession: AI2908
A>Status: preliminary
A.Molecule type: DNA
A.Residues: 1-437 <KUR>

[illegible]

Db 128 HHGGIINAIYVKGQVKEGDKLIELETRIKSEHENILGOYRNFPLATENRLAERDNL 187
QY 120 PSISPADLRSORADPVARAIADEQAQFTERROTIGQVDLMAQLOQOSEIEGIDRQ 179
Db 188 EQLEFSQFLMODINLPVAKIITHTQENLFRSKREYVSEKDALHONIAOLEKKEGLEAK 247
QY 180 TQGLKDLQLEFIEDELIDRLKYDKGLVPRRLALAEARSLSGSIGRLTADRKAVOGA 239
Db 248 KVAASKTAEVYQDRKALRTLKEKGFQKAALLDQEAQVAAKSDVATTEAETIGIRHAI 307
QY 240 SDPQAKRQIKQEFFEVSQSITEFVRLAEVTEKEVVAADQAKRITVSPVNGTAONLR 299
Db 308 TETQIKINQNNKTEETLTELREAOVOTASLEKYNALDLSLNRVITIRAPVDGIVSLK 367
QY 300 FTEGAVVRAAEPLVDIAPDEAFVIOAHFQPTVDVNVHMGVTEVRLPAFHSAGNP--- 356
Db 368 YHTIGVISHQGPIMEISPLNDPLIIIAKYSQKNIDSVHGLVAKIRFSAFKSTTTFET 427
QY 357 -DPERHDVAVADRISDP-QKQARLFLGIVRD-----VKQLRPHLRGRTVAGMPA 405
Db 428 GKAVYSIDPIVDERQYPGQQQDNVYVARVEIDMDEFNKVAKVKNLELH-----PGMQA 481
QY 406 QVIVPGERTVLOYFSPLDLTLTMTREE 435
Db 482 EVOIVGTGRTLRLADPVDITAFKAFREK 511

RESULT 12

D71687
alkaline proteinase secretion protein apre (apre) RP314 - Rickettsia prowazekii
C:Species: Rickettsia prowazekii
C>Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
C:Accession: D71687
R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Alsmark, U.
Nature 396, 133-140, 1998
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A:Reference number: A71630; MVID:99039499
A:Accession: D71687
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-511 <AND>
A:Cross-references: GB:AJ235271; GB:AJ235269; NID:93868717; PIDN:CAA14774.1; PID:9386087
A:Experimental source: strain Madrid E
C:Genetic:
A:Gene: apre; RP314

Query Match 24.08; Score 519.5; DB 2; Length 511;
Best Local Similarity 29.3%; Pred. No. 7.1e-25;
Matches 134; Conservative 93; Mismatches 192; Indels 39; Gaps 10;

QY 6 IORPTD---NFOAVAR--IGYGIALTPEVGLG--MAAFAPLDSAVIANGVSAEVS-Q 56
Db 65 IYTKTDNRNNVNAQAASPIIFGIYVIFLVLIGLWSALCPDSCGAVAGIVIPSTKK 124
QY 57 DVQHLBEGMLAKILVREGKVKAGOVLELDPQTQANNAAGITRNOYVALKAMEARLLAER 116
Db 125 TIQHNIEGIIINAIYVKGDKYKDEKDLIELETRIKSEHENILSQYNNPLATENRLAER 184
QY 117 DQRPISFPADLTSORAD-PMVARAIADDEQAQFTERROTIGQVDLMAQLOQOSEIEG 175
Db 185 DNLEQIKF-SDFLMQNNINLEPAKIIHTQENLFRSKREYVSEKDAFNQNTAOLEKIDG 243
QY 176 IDROTQGLKQQLGTELELDLRLKYDKGLVPRRLALAEARSLSGSIGRLTADRNSKA 235
Db 244 LEAKKIAASKTSEVYQDRKALRTLKEKGFQKAALLDQEAQVAAKSDVATTEAETIGI 303
QY 236 VQASDPTQAKRQIKQEFFEVSQSITEFVRLAEVTEKEVVAADQAKRITVSPVNGTA 295
Db 304 RHATTEQIKINQNNKTEETLTELREAOQVOTASLEKYNALDLSLNRVITIRSPVNGI 363
QY 296 QNREFTEGAVVRAAEPLVDIAPDEAFVIOAHFQPTVDVNVHMGVTEVRLPAFHSAGN 355

Db 364 NMLKYHTIGVISHQGPIMEISPLNDPLIIARIISQKNIDSVHGLVAKIRFSAFSRRT 423
QY 356 P-----DPERHDVAVADRISDPQKQARLFLGIVRD-----VKQLRPHLRG 397
Db 424 PTFPGKAVYSIDP-----IYDERQHILGQQQDNVYVARVEIDMDEFNKVAKVKNLALH--- 476
QY 398 RVTAGMPAOYIVPTGERTVLOYFSPLDLTLTMTREE 435
Db 477 ---PGMQAEVQIVGTGRTLRLADPVDITAFKAFREK 511

RESULT 13

S26697
apre protein - Pseudomonas aeruginosa
C:Species: Pseudomonas aeruginosa
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 31-Mar-2000
C:Accession: S26697
R:Duong, F.; Lazdunski, A.; Cami, B.; Murgier, M.
Gene 121, 47-54, 1992
A:Title: Sequence of a cluster of genes controlling synthesis and secretion of alkali
A:Reference number: S26696; MVID:93051361
A:Accession: S26697
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-432 <DO>
A:Cross-references: EMBL:X64558; NID:945279; PIDN:CAA45856.1; PID:945281
C:Superfamily: hemolysin secretion protein D; lipoyl/biotin-binding homology
C:Keywords: transmembrane protein
F:48-91,284-313/Domain: lipoyl/biotin-binding homology #status atypical <LPB>

Query Match 23.8%; Score 514.5; DB 2; Length 432;
Best Local Similarity 31.6%; Pred. No. 1.2e-24;
Matches 137; Conservative 88; Mismatches 182; Indels 27; Gaps 11;

QY 15 AVARIGGIALTFTVGLLGMARFAPLDS--AVIANGVSAEVSQDVQHLBEGMLAKILVR 72
Db 11 AYARLGMILVLEFGGALLMAAFAPLDQGVAVPATVLIISQ-RKSYQHPILGGVVKHILVR 69
QY 73 EGEKVKAGOVLELDPQTQANNAAGITRNOYVALKAMEARLLAERDQRPISPADLTSOR 132
Db 70 DQGVHEGEPLIRMEPEYQARNVDSLLNRVYANLQARLOAEYDGRRTLEMPAGLEQA 129
QY 133 ADPMVARAIADDEQAQFTERROT-IOGQVDLMAQLOQOSEIEGIDROTGS---LKDQLG 188
Db 130 PLPTLGBRL-ELQKQLHSQRTALANELSLRANIEGRLQLEGL-RQTEGNQRLQRL- 186
QY 189 FIEDELIDRLKYDKGLVPRRLALE-----ARAGSLSGSIGRLTADRKAVOGASD 241
Db 187 -LNSQLSGARDLAEQGYMPRQMLEQERQLAEVNARLSESSGRFGQIR-----QSIAT 238
QY 242 TQLKVRQIKQEFFEVSQSITEFVRLAEVTEKEVVAADQAKRITVSPVNGTAONLR 301
Db 239 AQMRITAAEETRYKREYVNGQLAETQVNAKRTLMEELSSARTELKHAETRAPSGIYAGIKVF 298
QY 302 TEGAVVRAAEPLVDIAPDEAFVIOAHFQPTVDVNVHMGVTEVRLPAFHSAGNPDEERH 361
Db 299 TDGQVIVPGBELMTIVVNSDSELEGGQAVLVDRIRHSGRLVLEMLFPAFNQSKPRVTGE 358
QY 362 DPAVAVADRISDPQKQARLFLGIVRVYDKQLRPHLRG-RVTAGMPAOYIVPTGERTVLOY 420
Db 359 VTWVSADRLDEQNKQKQYRVRAQVDAAM-GKLKGLQIRPGMAVQYFVTRGERSLNTYL 417
QY 421 FSPLDLTLTMTRE 434
Db 418 FKPLFDRAHVALAE 431

RESULT 14

G83489
alkaline proteinase secretion protein Apre PA1247 [imported] - Pseudomonas aeruginosa
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: G83489
C:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim, J.; Lory, S.; Olson, M.V.
N:Native 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337
A:Accession: G83489
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-432 <STO>
A:Cross-references: GB:AE004554; GB:AE004091; NID:99471174; PIDN:AG04636.1; GSPDB:GN001001
A:Experimental source: strain PA01
C:Genetics:
G:Gene: aprE; PA1247

| | | | | |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match | 23.8%; | Score 514.5; | DB 2; | Length 432; |
| Best Local Similarity | 31.6%; | Pred. No. 1.2e-24; | | |
| Matches 137; | Conservative 88; | Mismatches 182; | Indels 27; | Gaps 11; |

[illegible]

RESULT 15
S12526

C:Species: *Erwinia chrysanthemi*
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 26-May-2000
C:Accession: G12576.m86c57

C:\accessdb: 312320, 14000/
R:Letoffe, S.; Delepelalre, P.; Wandersman, C.

EMBO J. 9, 1375-1382, 1990

A:Reference number: S12524; MUID:90228333
A:Accession: S12526

A: Status: preliminary
A: Molecule type: DNA

A;Residues: 1-448 <LET>

A;Cross-references: GB:X53253; NID:g41546; PIDN:CAA37343.1; PID:g41549
R:Delepelatre, P.; Wandersman, C.

submitted to the EMBL Data Library, April 1991

A;Reference number: Z24521

```

A:Accession: T48667
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-448 <DEL>
A:Cross-references: EMBL:M60395; PIDN:AAA63635.1
A:Experimental source: strain B374
C:Genetics:
A:Gene: pTIE
C:Complex: the export system consists of the ATP-binding protein pTID (PIR:S12525); t
76141}
C:Function:
A:Description: involved in the export of metalloproteinases [validated, MUID:90228333
C:Superfamily: hemolysin secretion protein D; lipoyl/biotin-binding homology
C:Keywords: membrane-associated protein; protein export
C:64-107,300-329/Domain: lipoyl/biotin-binding homology #status atypical <LPB>

Query Match      23.4%   Score 506.5;   DB 2;   Length 448;
Best Local Similarity 28.8%;   Pred. No. 3.6e-24;
Matches 128; Conservative 93; Mismatches 191; Indels 33; Gaps 7;

```

| | | | | |
|-----------------------|------------------|--------------------|------------|-------------|
| Query Match | 23.4%; | Score 506.5; | DB 2; | Length 448; |
| Best Local Similarity | 28.8%; | Pred. No. 3.8e-24; | | |
| Matches 128; | Conservative 93; | Mismatches 191; | Indels 33; | Gaps 7; |

[illegible]

Search completed: July 19, 2002, 10:42:16
Job time: 76 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2002, 10:49:04 ; Search time 17.51 Seconds

(without alignments)
961.908 Million cell updates/sec

Title: US-09-913-414-5

Perfect score: 2165
Sequence: 1 MKPKIQHPDNTNFQAVARIG.....VLQYLSPDRDLRTTMRRE 435

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Query Length | DB ID | Description |
|------------|-------|-------|--------------|-------|-------------------------------|
| 1 | 514.5 | 23.8 | 432 | 1 | APRE_PSEAE 003025 pseudomonas |
| 2 | 506.5 | 23.4 | 448 | 1 | PRTE_ERWCH P23537 erwina chr |
| 3 | 313.5 | 14.5 | 477 | 1 | RT3D_ACTPL 008633 actinobacil |
| 4 | 288.5 | 13.3 | 478 | 1 | HLV4_ECOLI 009986 escherichia |
| 5 | 267.5 | 12.4 | 477 | 1 | HLVD_ACTAC P18790 actinobacil |
| 6 | 267.5 | 12.4 | 478 | 1 | HLVD_ECOLI P06739 escherichia |
| 7 | 266 | 12.3 | 440 | 1 | CYAD_BORPE P11091 bordetella |
| 8 | 243.5 | 11.2 | 478 | 1 | HLVD_PASPP P55125 pasteurella |
| 9 | 233.5 | 10.8 | 477 | 1 | RT1D_ACTPL P26761 actinobacil |
| 10 | 233 | 10.8 | 478 | 1 | HLVD_PASHA P16534 pasteurella |
| 11 | 164 | 7.6 | 424 | 1 | MCHE_ECOLI 09exn6 escherichia |
| 12 | 163 | 7.5 | 413 | 1 | CVAA_ECOLI P22519 escherichia |
| 13 | 147.5 | 6.8 | 343 | 1 | YUCH_ECOLI P32716 escherichia |
| 14 | 146 | 6.7 | 421 | 1 | EXOF_RHIME 002718 rhizobium m |
| 15 | 142 | 6.6 | 385 | 1 | YH1L_ECOLI P37626 escherichia |
| 16 | 136.5 | 6.3 | 352 | 1 | Y894_HAEIN 057500 haemophilus |
| 17 | 130.5 | 6.0 | 879 | 1 | CTPI_MYCTU 002171 onchocerca |
| 18 | 130 | 6.0 | 1625 | 1 | MYSP_ONCYO 010900 mycobacteri |
| 19 | 129.5 | 6.0 | 848 | 1 | MYSP_DIRIM P13382 dirofilaria |
| 20 | 128 | 5.9 | 371 | 1 | MACA_ECOLI P75830 escherichia |
| 21 | 126 | 5.8 | 882 | 1 | MYSP_CAEEL P10567 caenorhabdi |
| 22 | 126 | 5.8 | 390 | 1 | EMRA_ECOLI P27303 escherichia |
| 23 | 124 | 5.7 | 371 | 1 | MACA_YERPE P58411 yersinia pe |
| 24 | 124 | 5.7 | 880 | 1 | MYSP_BRUMA 001202 brugia mala |
| 25 | 123.5 | 5.7 | 1576 | 1 | RPOC_AOUPY 09x652 aquifex pyt |
| 26 | 122.5 | 5.7 | 378 | 1 | YIBH_ECOLI P32107 escherichia |
| 27 | 122 | 5.6 | 371 | 1 | MACA_ECO57 P58410 escherichia |
| 28 | 121.5 | 5.6 | 378 | 1 | YIAY_ECOLI P37683 escherichia |
| 29 | 121.5 | 5.6 | 545 | 1 | HTRS_HAUNI 048318 halobacteri |
| 30 | 121.5 | 5.6 | 4684 | 1 | PLEI_HUMAN 051549 homo sapien |
| 31 | 120 | 5.5 | 572 | 1 | YOG2_DEIRA 09rm6 delinococcus |
| 32 | 120 | 5.5 | 692 | 1 | MYL_POCCA 005000 podocoryne |
| 33 | 120 | 5.5 | 4473 | 1 | PLEI_CRIGR 09j155 cricetulus |

| | | | | | |
|----|-------|-----|------|---|-------------------------------|
| 34 | 119.5 | 5.5 | 474 | 1 | LCND_LACIA 000565 lactococcus |
| 35 | 118.5 | 5.5 | 440 | 1 | ODP2_ZYMO 066119 zymomonas m |
| 36 | 118 | 5.5 | 863 | 1 | MYSP_TAESO P35418 taenia soli |
| 37 | 117.5 | 5.4 | 1375 | 1 | RPOB_VIBCH 09kv30 vibrio chol |
| 38 | 117 | 5.4 | 652 | 1 | RPSD_CAVCR P52324 caulobacter |
| 39 | 116.5 | 5.4 | 449 | 1 | COMB_STRPN P36458 streptococc |
| 40 | 116 | 5.4 | 1238 | 1 | SBCC_RHOCA 068032 rhodobacter |
| 41 | 115 | 5.3 | 398 | 1 | ACOC_BACSU 031550 bacillus su |
| 42 | 115 | 5.3 | 995 | 1 | H1P1_HUMAN 000291 homo sapien |
| 43 | 114.5 | 5.3 | 387 | 1 | EMRK_ECOLI P52599 escherichia |
| 44 | 114.5 | 5.3 | 390 | 1 | EMRA_HAEIN P44928 haemophilus |
| 45 | 114.5 | 5.3 | 412 | 1 | MTRC_NEIGO P43505 neisseria g |

ALIGNMENTS

| RESULT | 1 | STANDARD | PRT | 432 AA. |
|------------|--|----------|-----|---------|
| APRE_PSEAE | | | | |
| ID | APRE_PSEAE | | | |
| AC | 003025: | | | |
| DT | 01-OCT-1993 (Rel. 27, Created) | | | |
| DT | 16-OCT-2001 (Rel. 40, Last sequence update) | | | |
| DT | 16-OCT-2001 (Rel. 40, Last annotation update) | | | |
| DE | Alkaline protease secretion protein apre. | | | |
| GN | APRE OR PA1247. | | | |
| OS | Pseudomonas aeruginosa. | | | |
| OC | Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae; | | | |
| OC | Pseudomonas. | | | |
| OX | NCBI_TaxID=287; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-ATCC 15692 / PA01; | | | |
| RX | MEDLINE=20437337; PubMed=10984043; | | | |
| RA | Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P., | | | |
| RA | Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M., | | | |
| RA | Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y., | | | |
| RA | Brody L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M., | | | |
| RA | Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T., | | | |
| RA | Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.; | | | |
| RT | "Complete genome sequence of Pseudomonas aeruginosa PA01, an | | | |
| RT | opportunistic pathogen."; | | | |
| RL | Nature 406:959-964(2000). | | | |
| CC | - FUNCTION: INVOLVED IN THE SECRETION OF ALKALINE PROTEASE. | | | |
| CC | - SUBCELLULAR LOCATION: INNER MEMBRANE-BOUND (POTENTIAL). | | | |
| CC | - SIMILARITY: BELONGS TO THE HLVD FAMILY OF SECRETION PROTEINS. | | | |
| CC | ----- | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ | | | |
| CC | or send an email to license@sib-sib.ch). | | | |
| CC | ----- | | | |
| DR | EMBL: X64588; CNA45856.1; - | | | |
| DR | EMBL: AE004554; AAG04636.1; - | | | |
| DR | PIR: S26697; S26697. | | | |
| DR | HSSP: P02905; IBD0. | | | |
| DR | InterPro: IPR002215; HLVD. | | | |
| DR | InterPro: IPR003997; RtxD. | | | |
| DR | Pfam: PF00529; HLVD; 1. | | | |
| DR | PRINTS: PR01490; RTXTOXIND. | | | |
| DR | PROSITE: PS00543; HLVD_FAMILY; 1. | | | |

| | |
|---|--|
| RC | STRAIN-SEROTYPE 2; PubMed=6494611. |
| RA | MEDLINE-93263992; Shin D.H.; |
| RA | Chang Y.-F., Shi J., Ma D.-P., Shin S.-J., |
| RT | "Molecular analysis of the Actinobacillus pleuropneumoniae RTX |
| RT | toxin-III gene cluster."; |
| RL | DNA Cell Biol. 12:351-362(1993). |
| RP | [2] |
| RP | SEQUENCE FROM N.A. |
| RC | STRAIN-405 / SEROTYPE 8; |
| RC | MEDLINE-95012630; PubMed=7927703; |
| RA | Jansen R., Briatore J., van Geel A.B.M., Kamp E.M., Gielkens A.L.J., |
| RA | Smits M.A.; |
| RT | "Genetic map of the Actinobacillus pleuropneumoniae RTX-toxin (Apx) |
| RT | operators: Characterization of the ApxIII operators."; |
| RL | Infect. Immun. 62:4411-4418(1994). |
| CC | -1- FUNCTION: INVOLVED IN THE TRANSPORT OF THE TOXIN RTX-III. |
| CC | -1- SUBCELLULAR LOCATION: INNER MEMBRANE-BOUND (POTENTIAL). |
| CC | -1- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF SEROTYPE 2. |
| CC | -1- SIMILARITY: BELONGS TO THE HLYD FAMILY OF SECRETION PROTEINS. |
| CC | ----- |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - |
| CC | the European Bioinformatics Institute. There are no restrictions on its |
| CC | use by non-profit institutions as long as its content is in no way |
| CC | modified and this statement is not removed. Usage by and for commercial |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ |
| CC | or send an email to license@isb-sib.ch). |
| CC | ----- |
| DR | EMBL: L12145; AAA21926.1; - |
| DR | EMBL: X80055; CAA56360.1; - |
| DR | InterPro: IPR002215; HLYD. |
| DR | InterPro: IPR003997; HLYD. |
| DR | Pfam: PF00529; HLYD. 1. |
| DR | PRINTS: PR01490; RTXTOXIND. |
| DR | PROSITE: PS00543; HLYD.FAMILY. 1 |
| KM | Cyclolysis: Transport; Transmembrane; Inner membrane. |
| FT | DOMAIN 1 59 |
| FT | TRANSMEM 60 77 |
| FT | DOMAIN 78 77 |
| FT | POTENTIAL. PERIPLASMIC (POTENTIAL). |
| FT | VARIANT 44 44 R -> A (IN SEROTYPE 8). |
| FT | VARIANT 132 132 C -> A (IN SEROTYPE 8). |
| FT | VARIANT 185 186 GA -> NT (IN SEROTYPE 8). |
| FT | VARIANT 376 376 V -> I (IN SEROTYPE 8). |
| SO | SEQUENCE 477 AA; 54780 MW; D2299FLCAB7D7E90A CRC64; |
| Query Match | 14.5%; Score 313.5; DB 1; Length 477; |
| Best Local Similarity | 25.3%; Pred. Num. 7.5e-13; |
| Matches 112; Conservative 88; Mismatches 190; Indels 53; Gaps 13; | |
| QY | 19 IGYGIALTFFGLGMAFAPLDSAVLNG-VSRAEVSODVHLEGMLAKLIVEGEGV 77 |
| DB | 60 IAYLMLFLPLAIY-ISIISKEIYASATGKLYVSGHSGKEIKPIFNALVKDIFVKGQGF 118 |
| QY | 78 KAGVLFEL-----DPTQANAAAGITRNQYVALKAMEARLLAERDQRPISF-PADLT 129 |
| DB | 119 EKGQLLNLFTALGCDADKQTKVSGLERLDGYRKSLYSI-EHNNPLDLEFNQADF 176 |
| QY | 130 S-QRADPVAVALADE-----QAQFERROTIGQVDLMAAQRLOYSOLEGIDR 178 |
| DB | 177 SVEQEDKTGAHLLITEQEFWOKQYKQELVYORQAKQTVLANIRKYESASRIE--- 232 |
| QY | 179 QTOGKLDQGLTEDELIDRLKLYDKGALVPRPRLLALNR-----AGSLGSGIGRLADRSK 234 |
| DB | 233 -----KELSDIKLKYDVKYSISKHELLAQENKRYEASSELVSYSHLEVESD 280 |
| QY | 235 AVAGASDTQLVROIKQEFPEVQSQSLTEFVRRLAEVKEKEVVASDAQRIK--IVSPVN 292 |
| DB | 281 LKAGQEDLKVLTQLEKSDILKELQGNIRKQNLTELEKNE-----QRLASTIRAPVS 334 |
| QY | 293 GTAQNLRFETGAVVRAEPLVLDIAPDEAFVIOAHFOPTVDVNHMGMTVEVRLPAFHS 352 |
| DB | 335 GTVOOLKTHRTGAVVTAETLMLVIAEDVDLEVSALIONKRVQGVFEVIOEAVIKETFPY 394 |

| | | | |
|----|--|---|-------------------------------------|
| Oy | 353 | ACNPDPERHDPVAVADRISDPQKARLEGLTVRVYKOLPPLRLG-RVTAGCPAQVYPT | 411 |
| Db | 395 | TRYGYLYKVKVITITLDAIEHPQ-LCLVFNSTIEINKKTLTDGDKIEILGSGMSVIAETKT | 453 |
| Oy | 412 | GERVLYOYLFSPRLTRTMR | 434 |
| Db | 454 | GRSVYISFLSPLESTESTESLRE | 476 |
| | 4 | RESULT | |
| | HLV4_ECOLI | STANDARD: | PRT: 478 AA. |
| AC | P09986 | HLV4_ECOLI | |
| DT | 01-MAR-1989 | (Rel. 10, Created) | |
| DT | 01-MAR-1989 | (Rel. 10, Last sequence update) | |
| DT | 16-OCT-2001 | (Rel. 40, Last annotation update) | |
| DE | Hemolysin secretion protein D, chromosomal. | | |
| OS | HLVD. | | |
| OS | Escherichia coli. | | |
| OC | Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; | | |
| OC | Escherichia. | | |
| OX | NCBI_TaxID=562; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | STRAIN=J96 / SEROTYPE O4; | | |
| RC | MEDLINE=85234404; Pubmed=3891743; | | |
| RA | Felmlée T., Pellett S., Welch R.A.; | | |
| RT | "Nucleotide sequence of an Escherichia coli chromosomal hemolysin"; | | |
| RL | J. Bacteriol. 163:94-105(1985). | | |
| RP | [2] | | |
| RP | TOPOLOGY. | | |
| RX | MEDLINE=92357011; Pubmed=1495479; | | |
| RA | Schnelein R., Gentschev I., Molienkopf H.-J., Goebel W.; | | |
| RT | "A topological model for the hemolysin translocator protein HLVD."; | | |
| RL | Mol. Gen. Genet. 234:155-163(1992). | | |
| CC | -1- FUNCTION: INVOLVED IN THE TRANSPORT OF HEMOLYSIN A. | | |
| CC | -1- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane. | | |
| CC | -1- SIMILARITY: BELONGS TO THE HLVD FAMILY OF SECRETION PROTEINS. | | |
| CC | ----- | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | |
| CC | use by non-profit institutions as long as its content is in no way | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ | | |
| CC | or send an email to license@isb-sib.ch). | | |
| CC | ----- | | |
| DR | EMBL; M10133; AAA23977.1; . | | |
| DR | PIR; D24433; LEECD | | |
| DR | InterPro; IPR002215; HLVD. | | |
| DR | InterPro; IPR003997; RCLD. | | |
| DR | Pfam; PF00529; HLVD. 1. | | |
| DR | PRINTS; PRO1490; RTYTOXIND. | | |
| DR | PROSITE; PS00543; HLVD_FAMILY. 1. | | |
| KW | Hemolysis; Transport; Transmembrane; Inner membrane; Signal-anchor. | | |
| FT | DOMAIN 1 59 | CYTOPLASMIC (PROBABLE). | |
| FT | TRANSMEM 60 80 | SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN) | |
| FT | FT | (PROBABLE). | |
| FT | PERIPLASMIC (PROBABLE). | | |
| SO | SEQUENCE 478 AA; 54591 MW; 9E9ED0C42EC8D8089 CRC64; | | |
| | Query Match | 13.3%; | Score 288.5; DB 1; Length 478; |
| | Best Local Similarity | 25.2%; | Pred. No. 2.8e-11; |
| | Matches 116; Conservative | 94; | Mismatches 192; Indels 59; Gaps 18; |
| Oy | 1 | MKPPPTORPTNQFQAVARIGYIALTPV-GILGMAAFAPLDSANVANGVVS-AEVSQDV | 58 |
| Db | 49 | IETPVSRKP-----RIVAATFMGLVIAIILSLVG-----QVEIYATANGKLILSRSKET | 99 |
| Oy | 59 | QHLDEGMAKILIVREGCEVVKAGOVLFELDPTQANAAAGITNQYVALAA-----MEARLLA | 114 |

```
Db 100 KPIENSIIVEIKESVYRKGVLLKL--TALGAEDTLKQSSLSLQARLEIRQIILS 157
OY 115 ---EPDQPSISPADLTSSOR-ADPMVARAIADQAF--TERQQTQGGYDL--MNAQRL 167
Db 158 RSTELNKLPLKLPDEYKFPNVEEELRLSLIKEDFSFWQMOQKOKELNLDKKRERL 217
OY 168 QYSEIEGIDRQGTQKLDQGLFEDELIDRLKLYDKGLVPRPLLAEAR---AGSLG 223
Db 218 TILAIRIRYENVSREKSRDL-----DPSRLIKQAIKAKHVLQEDENKYEVAANLRY 270
OY 224 SISRRLADRSKAVQASDPLQKROIKEFFEQYSSITFRVLAVTEKEVYASDAQ 283
Db 271 YKSOLEIESEILSAKEEYQLVQLFKNELDLKQRTTDSITELTLEKNE---ERQQ 326
OY 284 RIRIVSPVNGTAQNLRFETEGAVVRAEPLVIDAPEDAEVIAQHPOPTDVNVMGMVT 343
Db 327 ASVIRAPVSKVQGLKHTGGVVTATLMLVPEDDTLEVTALYQNKDIGFINVQNA 386
OY 344 EVRLPAFHSAGNPDPERHDPVANA-----DRISDQKQARLFLGIVRDYKOLP--P 393
Db 387 IIRVEAF-----PYTRYGYLVGKVKKNIMLDAIED--QKGLVFENVIVSVEENDLSTGNK 438
OY 394 HLRGRTVAGMPAQIYPTGERTVQLYLFSPLRDLRTTMBE 434
Db 439 HI--PLSSGMAVTAETIKTGMRSVSYLSPLESVTESLHE 477

RESULT 5
HLVD_ACTAC STANDARD; PRT; 477 AA.
AC P18790;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Leukotoxin secretion protein D.
GN LKTD OR AALTD.
OS Actinobacillus actinomycetemcomitans (Haemophilus
OC actinomycetemcomitans).
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Actinobacillus.
OX NCBI_TaxID=714;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JP2;
RA MEDLINE=90384843; Pubmed=2402458;
RA Guttmiller J.M., Kraig E., Cagle M.P., Kolodrubetz D.;
RT "Sequence of the lktD gene from Actinobacillus
RT actinomycetemcomitans."
RL Nucleic Acids Res. 18:5292-5292(1990).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=92072446; Pubmed=1961107;
RA Lally E.T., Golub E.E., Kleba I.R., Taichman N.S., Decker S.,
RA Berthold P., Gibson C.W., Demuth D.R., Rosenblom J.;
RT "Structure and function of the B and D genes of the Actinobacillus
RT actinomycetemcomitans leukotoxin complex."
RL Microb. Pathog. 11:111-121(1991).
CC -1- FUNCTION: INVOLVED IN THE TRANSPORT OF THE HEMOLYSIN/LEUKOTOXIN.
CC -1- SUBCELLULAR LOCATION: INNER MEMBRANE-BOUND (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE HLVD FAMILY OF SECRETION PROTEINS.
CC -1- SIMILARITY: THE N-TERMINAL (AA 1-153) IS SIMILAR TO E.FAECALIS
CC PORE FORMING PROTEIN EBSA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X53956; CA37907.1; -
DR PIR: S11215; S11215.
```

```
DR PIR: B61378; B61378.
DR InterPro: IPR002215; HLVD.
DR InterPro: IPR003997; RtxD.
DR Pfam: PF00529; HLVD; 1.
DR PRINTS: PR01490; RTXQXIND.
DR PROSITE: PS00543; HLVD_FAMILY; 1.
KW Hemolysin; Cytolysis; Transport; Transmembrane; Inner membrane.
FT DOMAIN 1 59
FT TRANSMEM 60 80
FT DOMAIN 81 477
FT POTENTIAL.
SQ SEQUENCE 477 AA; 54651 MW; 54F20128CADE6260 CRC64;

Query Match 12.4%; Score 267.5; DB 1; Length 477;
Best Local Similarity 23.7%; Pred. No. 5.9e-10;
Matches 108; Conservative 87; Mismatches 207; Indels 53; Gaps 15;

OY 6 IORTDNEQAVARIGVGIILTFVGLGMAPPLDVAIVANGVS-AEVSODVQHLGG 64
Db 49 IETPVSN--APREVSYIMLFLLAI--VSIFSVEIATASGKFKALSGRSKEIKPIENS 105
OY 65 MLAKIIVREGEKVACGVLEFLDPTQANAAAGITRNQVALKAMEARL-----LAERDQR 119
Db 106 LVKHIFVKEGEYVKKGELLKLALGAEADTLKKTSLQAKLEEFYKSLLEAVERDQL 165
OY 120 PSISF-PADLTSSRADPMVARAIADQAF--TERQQTQGGYDL--MNAQRLQYSEIEGI 176
Db 166 PILDFSKIDLPFTMNQKRVTLIEQFSTWQKQKQKTLNKKDAEKLSTYARL--- 222
OY 177 DRQQTQKLDQGLFEDE---LIDRLKLYDKGLVPRPLLLEARAGSLSSIGRTADRS 233
Db 223 -----KTYEGLINTQVRLDDEFRALYKEHAIKHTVLDEE---NKYQAINLEIYVKA 272
OY 234 KAVQ-----GASDTQLKVKROIKEFFEQYSSITFRVLAEVTEKEVYASDAQRIK 286
Db 273 SLQGVENEVLAKEEQELVQGLFKNDILDKLQKTDVNNILTFELDKN---NQKQVSE 328
OY 287 IVSPVNGTAQNLRFETEGAVVRAEPLVIDAPEDAEVIAQHPOPTDVNVMGMVTEVR 346
Db 329 IRAPVSGVQDLKVTHTIDGVVTTAETLMVAVVEDSLEVTALLQNKDIGVKEQOEYVIR 388
OY 347 LPAF-----HSAGNPDPERHDPVANAADRISDQKQARLFLGIVRDYKOLPRLRG--RV 399
Db 389 VEAFPPYTRYGYLTGKVKV-----ITLDAIHP--KGLGVFTITIEDKLTSTEEKETPL 441
OY 400 TAGMPAQIYPTGERTVQLYLFSPLRDLRTTMBE 434
Db 442 SAGMEITAEIKTGMRSVSYLSPLESIDKSLRE 476

RESULT 6
HLVD_ECOLI STANDARD; PRT; 478 AA.
AC P06739;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hemolysin secretion protein D, plasmid.
GN HLVD.
OS Escherichia coli.
OC Plasmid Incir pHLV152.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RA Hess J., Wels W., Vogel M., Goebel W.;
RT "Nucleotide sequence of a plasmid-encoded hemolysin determinant and
RT its comparison with a corresponding chromosomal hemolysin sequence."
RL FEMS Microbiol. Lett. 34:1-11(1986).
RN [2]
RP TOPOLOGY
RA MEDLINE=92357011; Pubmed=1495479;
```



```

Db 376 VDDROQHSYRVITIALHAPALEVDGK---PRL---LKEGNAVQADIRTSRLLIEVLLSPV 429

RESULT 8
ID HLVD_PASSP STANDARD: PRT; 478 AA.
AC P53125;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Leukotoxin secretion protein D.
GN LKTD.
OS Pasteurella haemolytica-like sp. (strain 5943B).
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OC NCBI_Taxid=28165;
OX NCBI_Taxid=28165;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=93239320; PubMed=8478098;
RX Chang Y.-F., Ma D.-P., Shi J., Chengappa M.M.;
RA "Molecular characterization of a leukotoxin gene from a Pasteurella
RA haemolytica-like organism, encoding a new member of the RTX toxin
RA family."
RT Infect. Immun. 61:2089-2095(1993).
RL -1- FUNCTION: INVOLVED IN THE TRANSPORT OF THE LEUKOTOXIN.
CC -1- SUBCELLULAR LOCATION: INNER MEMBRANE-BOUND (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE HLVD FAMILY OF SECRETION PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L12148; AAA16446.1; -
CC InterPro: IPR002215; HLVD.
CC DR InterPro: IPR003997; RTX.D.
CC Pfam: PF00529; HLVD; 1.
CC PRINTS: PR01490; RTXTOXIND.
CC DR PROSITE: PS00543; HLVD_FAMILY; 1.
CC KW Hemolysis; Cytolysis; Transport; Transmembrane; Inner membrane.
CC FT DOMAIN 1 77 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 78 98 POTENTIAL.
CC FT DOMAIN 99 478 PERIPLASMIC (POTENTIAL).
CC SO SEQUENCE 478 AA; 54650 MW; 221461A69482913A CRC64;

Query Match 11.2%; Score 243.5; DB 1; Length 478;
Best Local Similarity 24.4%; Pred. No. 1.9e-08;
Matches 113; Conservative 73; Mismatches 174; Indels 103; Gaps 18;

QY 31 LLGNAAPPLDSDAVIANGVSAEV-----SQDVHLEGGLAKITLVREGKV 77
DB 59 LIAVLIMFLAVLAVLASVSEVEIATAPAGRLTFSGRSKETPTETVIOEIFVKGQGV 118
QY 78 KAGOVLEFLDP--TGANAAGITRNQVYALAKAMEARLL--AERDORSISFPADLT--- 129
DB 119 EKGQLVSLTALGSDADIKTITISLAKENRYQTLLTIEKESLPVT---DLSTTE 174
QY 130 ---SORADPMVARIADP-----QAOFTER-----ROTTIOGVYDLMNAORLOYOS 171
DB 175 FKDSSEDRRLKIKHLEIQYTTWOKOTKLAVKRKAERQOTISA-----YVR 223
QY 172 EIEGIDROTQGIKQDLGTEDELIDRLKLYKGLVPRRLALLARAGSLGSGIGRLTAD 231
DB 224 KYEGATRIEQ-----EKLDFKRLYQOKSLSKHELLSOENKALEAO---NELAVY 270
QY 232 RSKAVOGASD-----TQLKVRQIKOEFPOVSOSITETPVRLAVEKEVVASDAOKR 284
DB 271 RSKLNEESDLINKELELITQFFKSVDLEKRLKOHINERQOLEL-EKN---NORQOA 326

QY 285 IKIVSPVNGTAQNLKRFTEGAVYRAAEPLVDIAPDEAEVIOAHPOPTDVNVHMGVTE 344
DB 327 SMIRAPVSGTVOQLKTHITIGVTTAETLMTIVPEEDVLEATLQNKDIGFPAAGOEVI 386
QY 345 VRLPAPFASAGNPDPERRHDPVA-----VADRIS-----DPOKARL---FLGIVRDVQL-- 391
DB 387 IKVETP-----PYTRGYITGRKIKHSPALIEOPNGLGVFNATVSDIKQALSS 434
QY 392 PPHLRGRVTAGMPAOYIVPTGERTVLOYLSPRLDRLRTTMR 434
DB 435 PDGKHTELGGMITTAETKIGERSVMSYLLSPLEESVTESLRE 477

RESULT 9
RTID_ACTPL RTID_ACTPL STANDARD: PRT; 477 AA.
AC P26761;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE RTX-I toxin determinant D (APX-ID) (HLV-ID) (Cytolysin ID) (CLY-ID)
DE (Toxin RTX-I secretion protein D).
GN APXID OR CLYID OR HLVD OR APPD.
OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Actinobacillus.
OC NCBI_Taxid=715;
OX NCBI_Taxid=715;
RN [1]
RN SEQUENCE FROM N.A.
RX STRAIN=SEROTYPE 5;
RX MEDLINE=91317735; PubMed=1860823;
RX Chang Y.-F., Young R., Struck D.K.;
RA "The Actinobacillus pleuropneumoniae hemolysin determinant: unlinked
RA appca and appbd loci flanked by pseudogenes."
RT J. Bacteriol. 173:5151-5158(1991).
RL [2]
RN SEQUENCE FROM N.A.
RX STRAIN=ISOLATE CVI 13261 / SEROTYPE 9;
RX MEDLINE=92040145; PubMed=1937809;
RA Smits M.A., Brilaire J., Jansen R., Smith H.E., Kamp E.M.,
RA Gielkens A.L.;
RT "Cytolysins of Actinobacillus pleuropneumoniae serotype 9."
RL Infect. Immun. 59:4497-4504(1991).
RN [3]
RN SEQUENCE FROM N.A.
RX STRAIN=S 4074 / SEROTYPE 1;
RX MEDLINE=94237497; PubMed=8181764;
RA Frey J., Haldemann A., Nicolet J., Boffini A., Prentki P.;
RT "Sequence analysis and transcription of the apxi operon (hemolysin I)
RT from Actinobacillus pleuropneumoniae."
CC CC -1- FUNCTION: INVOLVED IN THE TRANSPORT OF THE TOXIN RTX-I AS WELL AS
CC THAT OF RTX-II.
CC -1- SUBCELLULAR LOCATION: INNER MEMBRANE-BOUND (POTENTIAL).
CC -1- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF SEROTYPE 5.
CC -1- SIMILARITY: BELONGS TO THE HLVD FAMILY OF SECRETION PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M65808; AAB00967.1; -
CC EMBL; X61112; CAA43426.1; -
CC DR EMBL; X68595; CAA48588.1; -
CC DR PIR; B40366; B40366.
CC DR PIR; S18856; S18856.
CC DR InterPro: IPR002215; HLVD.
CC DR InterPro: IPR003997; RTX.D.
CC Pfam; PF00529; HLVD; 1.

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2002, 10:48:41 ; Search time 53.06 Seconds
(without alignments)
1418.258 Million cell updates/sec

Title: US-09-913-414-5
Perfect score: 2165
Sequence: 1 MKPKIQRPTDNFOAVARIG.....VLQYLSPDLRLRTTMRRE 435

Scoring table:
BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SPREMBL_19:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*
16: sp_virus:*
17: sp_bacteriap:*
sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1 | 1984.5 | 91.7 | 436 | 16 | 085351 |
| 2 | 1972.5 | 91.1 | 436 | 2 | Q9RMM9 |
| 3 | 606 | 28.0 | 436 | 16 | Q92NXX |
| 4 | 589 | 27.2 | 436 | 2 | Q9X6N7 |
| 5 | 583 | 26.9 | 433 | 16 | Q33679 |
| 6 | 564 | 26.1 | 443 | 16 | Q9HYJ9 |
| 7 | 560.5 | 25.9 | 434 | 16 | Q98LQ9 |
| 8 | 543.5 | 25.1 | 473 | 16 | P96438 |
| 9 | 540.5 | 25.0 | 443 | 2 | Q54417 |
| 10 | 536.5 | 24.8 | 473 | 2 | Q05198 |
| 11 | 530.5 | 24.5 | 511 | 16 | Q921J2 |
| 12 | 522 | 24.1 | 443 | 2 | Q9XB63 |
| 13 | 521.5 | 24.1 | 443 | 2 | Q54457 |
| 14 | 519.5 | 24.0 | 511 | 16 | Q92DL4 |
| 15 | 512 | 23.6 | 435 | 2 | Q05694 |
| 16 | 510 | 23.6 | 433 | 2 | Q9ZG93 |

| | | | | | | |
|----|-------|------|-----|----|--------|---------------------|
| 17 | 509 | 23.5 | 437 | 2 | 067994 | 067994 pseudomonas |
| 18 | 502.5 | 23.2 | 447 | 16 | Q988E5 | Q988E5 rhizobium 1 |
| 19 | 499 | 23.0 | 435 | 2 | Q9FCN8 | Q9FCN8 rhizobium 1 |
| 20 | 495 | 22.9 | 444 | 2 | Q9ZNT8 | Q9ZNT8 pseudomonas |
| 21 | 493 | 22.6 | 437 | 2 | 087809 | 087809 pseudomonas |
| 22 | 489.5 | 22.6 | 427 | 2 | Q921G7 | Q921G7 campylobact |
| 23 | 482.5 | 22.3 | 427 | 2 | Q9R8D9 | Q9R8D9 campylobact |
| 24 | 479.5 | 22.1 | 443 | 2 | Q9RHT1 | Q9RHT1 pseudomonas |
| 25 | 479 | 22.1 | 439 | 2 | Q9R9H4 | Q9R9H4 pseudomonas |
| 26 | 477 | 22.0 | 438 | 2 | Q9KCS5 | Q9KCS5 pseudomonas |
| 27 | 476.5 | 22.0 | 450 | 16 | Q9ZLS8 | Q9ZLS8 rhizobium m |
| 28 | 471.5 | 21.8 | 437 | 2 | Q57387 | Q57387 serratia ma |
| 29 | 467.5 | 21.6 | 440 | 16 | Q98140 | Q98140 rhizobium 1 |
| 30 | 467.5 | 21.6 | 441 | 2 | 085376 | 085376 proteus mir |
| 31 | 436 | 20.1 | 433 | 2 | Q52859 | Q52859 rhizobium 1 |
| 32 | 368 | 17.0 | 481 | 16 | Q9KRM3 | Q9KRM3 vibrio chol |
| 33 | 367 | 17.0 | 335 | 16 | Q912M0 | Q912M0 pseudomonas |
| 34 | 329 | 15.2 | 475 | 16 | Q9JY53 | Q9JY53 neisseria m |
| 35 | 326.5 | 15.1 | 475 | 2 | Q9X580 | Q9X580 neisseria m |
| 36 | 326.5 | 15.1 | 475 | 2 | Q9JPK4 | Q9JPK4 neisseria m |
| 37 | 314 | 14.5 | 473 | 16 | Q9PAU7 | Q9PAU7 xyloella fas |
| 38 | 288.5 | 13.3 | 464 | 2 | P74828 | P74828 sphingomona |
| 39 | 283.5 | 13.1 | 439 | 2 | 087505 | 087505 escherichia |
| 40 | 276.5 | 12.8 | 478 | 2 | 008269 | 008269 escherichia |
| 41 | 272.5 | 12.6 | 468 | 16 | Q98MG5 | Q98MG5 rhizobium 1 |
| 42 | 271.5 | 12.5 | 477 | 2 | Q9RCG6 | Q9RCG6 pasteurella |
| 43 | 264 | 12.2 | 438 | 2 | 068080 | 068080 rhodobacter |
| 44 | 261.5 | 12.1 | 454 | 16 | Q9JTK4 | Q9JTK4 neisseria m |
| 45 | 251.5 | 11.6 | 479 | 2 | 082886 | 082886 escherichia |

ALIGNMENTS

| RESULT | ID | PRELIMINARY | PRT | 436 AA. |
|--------|---|-------------|------------------------|---------|
| 085351 | 085351 | | | |
| AC | 085351 | | | |
| DT | 01-NOV-1998 (TREMBLE) | 08 | Created | |
| DT | 01-MAY-2000 (TREMBLE) | 13 | Last sequence update | |
| DT | 01-DEC-2001 (TREMBLE) | 19 | Last annotation update | |
| DE | MEMBRANE FORMING UNIT (RSAA SECRETION SYSTEM, MEMBRANE PROTEIN | | | |
| DE | RSAA). | | | |
| GN | RSAA OR CC1009. | | | |
| OS | Caulobacter crescentus. | | | |
| OC | Bacteria; Proteobacteria; alpha subdivision; Caulobacter group; | | | |
| OC | Caulobacter. | | | |
| OX | NCBI_TaxID=69394; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-ATCC 19089 / CH15, AND NA1000; | | | |
| RC | AWRAM P., Smit J.K.; | | | |
| RX | MEDLINE-98292737; PubMed-9620934; | | | |
| RA | AWRAM P., Smit J.K.; | | | |
| RT | "The Caulobacter crescentus paracrystalline S-layer protein is | | | |
| RT | secreted by an ABC transporter (type I) secretion apparatus."; | | | |
| RL | J. Bacteriol. 180:3062-3069(1998). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-ATCC 19089 / CH15, AND NA1000; | | | |
| RC | AWRAM P., Smit J.K.; | | | |
| RA | Submitted (AFR-1998) to the EMBL/GenBank/DBJ databases. | | | |
| RN | [3] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-ATCC 19089 / CH15, AND NA1000; | | | |
| RC | AWRAM P.; | | | |
| RA | Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases. | | | |
| RN | [4] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-ATCC 19089 / CH15; | | | |
| RC | MEDLINE-21173698; PubMed-11259647; | | | |
| RX | Nierman W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E., | | | |
| RA | Eisen J., Heidelberg J.F., Alley M.R., Ohta N., Maddock J.R., | | | |
| RA | Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B., | | | |

RA DeBooy R.T., Dodson R.J., Durrkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Barry K.,
RA Utechtack T., Tian K., Wolf A., Yamathuan J., Eniolaaya M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.,
RT "Complete genome sequence of *Caulobacter crescentus*,"
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AF062345; AAC38667.2;
DR EMBL; AE005779; AAC2293.1; -;
DR TIGR; CCI009; -;
DR InterPro: IPR002215; HLYD.
DR InterPro: IPR003997; RtxD.
DR Pfam: PF00529; HLYD; 1.
DR PRINTS: PR01490; RYTXOIND.
DR PROSITE: PS00543; HLYD_FAMILY; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 436 AA; 48389 MW; 90E32A834673C9DC CRC64;

| | | | | |
|---------------------------|--------|---------------------|-----------|-------------|
| Query Match | 91.7%; | Score 1984.5; | DB 16; | Length 436; |
| Best Local Similarity | 93.6%; | Pred. No. 1.2e-117; | | |
| Matches 408; Conservative | 3; | Mismatches 24; | Indels 1; | Gaps 1; |

| | | | |
|----|-----|--|-----|
| Qy | 1 | MKPKRIORPDMPOAVARIGGIIATLFEVGLGMAAPAPLDSAVINGVSAFVS -ODVQ | 59 |
| | | : | |
| Dd | 1 | MKPPRIQPIPDNFOAVARIGGIIATLFEVGLGMAAPAPLDSAVIANGVSAEGNKRTYQ | 60 |
| Qy | 60 | HLEGGMLAKLIVRGEKVKAGOVLFELDPTQAAAAGITRNOYVALKAMEARLLAERDOR | 119 |
| | | | |
| Dd | 61 | HLEGGMLAKLIVRGEKVKAGOVLFELDPTQAAAAGITRNOYVALKAMEARLLAERDOR | 120 |
| Qy | 120 | PSISPRADLTQSQRADPMYARAIADEQAQOFERRCOTOGQVDMNAORLOYQOSEIBIDNQ | 179 |
| | | | |
| Dd | 121 | PSISPRADLTQSQRADPMYARAIADEQAQOFERRCOTOGQVDMNAOQLOQOSEIBIDNQ | 180 |
| Qy | 180 | TOGLKQDQGFIEDELDLIRKLUDYDKGLVPRRLTLALEARAGSLSGISGRLLADRSKAAVOGA | 239 |
| | | | |
| Dd | 181 | TOGLKQDQGFIEDELDLIRKLUDYDKGLVPRRLTLALEARAGSLSGISGRLLADRSKAAVOGA | 240 |
| Qy | 240 | SDTOLKQVQIKOEFPPEVOSOSITETRYRLAEVTEKEVYASDAQRIKITYSPVNGTQONLR | 299 |
| | | | |
| Dd | 241 | SDTOLKQVQIKOEFPPEVOSOSITETRYRLAEVTEKEVYASDAQRIKITYSPVNGTQONLR | 300 |
| Qy | 300 | PFTTEGAUVRAAEPLVDIAPEDEAEFVIOAHFOPTVDVNVHMGMTYEVLPAFHSAGNPDE | 359 |
| | | | |
| Dd | 301 | PFTTEGAUVRAAEPLVDIAPEDEAEFVIOAHFOPTVDVNVHMGMTYEVLPAFHSREIRPLN | 360 |
| Qy | 360 | RHDPVAVADRISDPQOARLEFVIGVRVDVKOLPRHLRGVATAGNPAOVIVPTGERTVLOY | 419 |
| | | : | |
| Dd | 361 | GTIQSLSDRISDRISDPQNKLDYETFLGIVRVDKOLPRHLRGVATAGNPAOVIVPTGERTVLOY | 420 |
| Qy | 420 | LFSPLADLTLTMTMREE 435 | |
| | | | |
| Dd | 421 | LFSPLADLTLTMTMREE 436 | |

RESULT 2

| | | | | | |
|----|-------|--------------|------|-----|-----|
| ID | Q9RM6 | PRELIMINARY; | PRT; | 436 | AA. |
|----|-------|--------------|------|-----|-----|

| | |
|----|-------------------------------------|
| DT | 01-MAY-2000 (Tremblay, 13, Created) |
| DT | 01-MAY-2000 (Tremblay, 13, Created) |

DT 01-JUN-2001 (TREMBLER, I/, Last annotation update)

GN KSAE. *Caulobacter crescentius*

Caulobacter.

RN [1]

RC STRAIN=JS4000;

RT "The secretion signal of the *Caulobacter crescentus* S-layer protein is

RT located within the C-terminal 82 amino acids of the molecule.²
 RL Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF193064; AAF67963.1; -
 DR InterPro: IPR002215; HLYD.
 DR InterPro: IPR003897; RCD.
 DR Pfam, PF00529; HLYD; 1
 DR Pfam, PRO1490; RYTOXIND
 DR PRINTS: P500543; HLYD.FAMILY:
 DR PROSITE: P500543; HLYD.FAMILY: UNKNOWN_1
 SQ SEQUENCE 436 AA; 48443 MW; F3C4C6B0F3D96563 CMC6;

| | | | | |
|------------------------------|--------|----------------------|-------|-------------|
| Query Match | 91.18; | Score 1972.5; | DB 2; | Length 436; |
| Best Local Similarity | 93.18; | Pred. No. 6.7e-117; | | |
| Number of Misclassifications | 3; | W/misclassifications | 26; | Indels 1; |
| | | | | Gaps 1 |

| | | | |
|----|-----|---|-----|
| Qy | 1 | MKPPRTOEPNPOAVARIGIGIALTFFVGLGMAAPAPLDSAVIANGVNSLEVS - ODVO | 59 |
| Qy | 1 | MKPPRTOEPNPOAVARIGIGIALTFFVGLGMAAPAPLDSAVIANGVNSLEVS - ODVO | 60 |
| Db | 1 | MKPPRTOEPNPOAVARIGIGIALTFFVGLGMAAPAPLDSAVIANGVNSLEVS - ODVO | 60 |
| Qy | 60 | HLEGGMLAKILVREGKVKAGVLELDPTOANAAAGITNNOYVALKAMBARLLAERDOR | 119 |
| Db | 61 | HLEGGMLAKILVREGKVKAGVLELDPTOANAAAGITNNOYVALKAMBARLLAERDOR | 120 |
| Qy | 120 | PSISPRALUTSQRADPMYARIADEQAOFTERRTOTOGOVUDIMNAORLOYOSEIEGIDQO | 179 |
| Db | 121 | PSISPRALUTSQRADPMYARIADEQAOFTERRTOTOGOVUDIMNAORLOYOSEIEGIDQO | 180 |
| Qy | 180 | TGGLDLDOLGFEDELELDLRLKLYDKGLVPRRLYLLAERAGSLSGSIGRLTADRSKAVOGA | 239 |
| Db | 181 | TGGLDLDOLGFEDELELDLRLKLYDKGLVPRRLYLLAERAGSLSGSIGRLTADRSKAVOGA | 240 |
| Qy | 240 | SPTOLKRVKIOKEPFEVOSISTETRVRLAEVTEKEVYASDAQRIKITYSPVNGTQONLR | 299 |
| Db | 241 | SPTOLKRVKIOKEPFEVOSISTETRVRLAEVTEKEVYASDAQRIKITYSPVNGTQONLR | 300 |
| Qy | 300 | FPTGEGVYRAAPRLVDIARBEDAEFYIOAHFQPTVDVNMHGMQVTEVRLPAFHSAGNDPE | 359 |
| Db | 301 | FPTGEGVYRAAPRLVDIARBEDAEFYIOAHFQPTVDVNMHGMQVTEVRLPAFHSAGNDPE | 360 |
| Qy | 360 | RHDPAVADRISSPOQARLEFLGIVRVUDKOLPRHLKGRVTAGMRAOVIVPTGERTVLOY | 419 |
| Db | 361 | GFTQISLODRISDPONKLDYELFGIARVDVKOLPRHLKGRVTAGMRAOVIVPTGERTVLOY | 420 |
| Qy | 420 | LFSPRLDRLRTTMREE 435 | |
| Db | 421 | LFSPRLDRLRTTMREE 436 | |

RESULT 3

| | | | |
|--------|--------------|------|---------|
| ID | PRELIMINARY; | PRT; | 436 AA. |
| Q92NX2 | | | |

DT 01-DEC

01-DEC-2001 (ITEMBLER: 19, Last annotation update)

GN Rhizobium meliloti (sinorhizobium meliloti).
SMC04208.

Rhizobiaceae; Sinorhizobium.

| | |
|----|-------|
| 00 | _____ |
| RN | [1] |

RC STRAIN=1021;

Galibert F., Finan T.M., Long S.R.

| | | | | | |
|----|------------|-----------|---------------|-----------------|-------------|
| RA | Cowie A | Davis P W | Dreano S | Federspiel N.A. | Fisher R.F. |
| KA | Boultry M. | Bowser L. | Bullmester J. | Caicedo E. | Capera D. |

RA Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,

RA Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B.,

RA Ransperger U., Surzycki R., Thebault P., Vandenberg M.,
RA Vorholzer F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;
RT "The composite genome of the legume symbiont *Sinorhizobium meliloti*.";
RL Science 293:668-672(2001).
DR EMBL: AL591789; CAC46615.1; -.
KW Complete proteome.
SQ SEQUENCE 436 AA; 47877 MW; 31E8A8F3180C872 CRC64;

Query Match 28.0%; Score 606; DB 16; Length 436;
Best Local Similarity 34.0%; Pred. No. 1,5e-30;
Matches 145; Conservative 75; Mismatches 180; Indels 26; Gaps 6;

QY 24 IALFVGLG-WAFAPLDSAVIANG--VSAEVSODVOHEGMLAKILVEGEGKAG 80
D 23 VGLALIGAGMATTTQLAGAVASGNFYVSHVK-VQHTGVIIGTILVEEGDWKAG 81
QY 81 QVLFELPTQANAAAGITRNQYALKAMEARLAEKDQPSISFPADITSGADPMVAA 140
D 82 QVLRPLDPTQANLRIYTKRLDELMLARLAEARDQGELEFTELKAKRAVASA 141
QY 141 IADQADFTERTOTGOVDLMNQRLQOSEIGIDROTGLDQGLFIDELIDRLK 200
D 142 INGSKLFEXRRRSRREGRSQLEIRIVQLHEIAGLDAQQAYDRGSLIDAEIASL 201
QY 201 YDKGLVPRPLLEARAGSLSGISGLTLADRSKAVQASDTOLKVRQKQEFEEVSOS 260
D 202 HERIVSVQRNLETRNATPEGLRGEKIAQQAAGRIATRIQIITQIDEDLKTENG 261
QY 261 ITRFRVLAETKEVVAASDAQRIKIVSPVNGTAQNLRFETEGAVVAAPLVDIAD 320
D 262 LREVOALGEPTEKRIAEQDLRRIDMLAPQGIHVHQLNHTVGVISPADVMSIYVDS 321
QY 321 EAFIQAHPOTDVNHHMGVTEVRLPAFHSAGNPDERIDPAVAVDRIDPQKARLF 380
D 322 DRLAIEKIRPQDIDQIQLGDDVLRLSAFNQRTTPELGGSVSRIADLTEDPSGLSY 381
QY 381 LGIVRVN-----KQPLHLGRVYAGPAQIVYGERIVQLYFSPLRDLR 429
D 382 --VVRVAVPGAEELARLSLELP-----GMPAEFIOTGORTALSTYIKLSQIS 430
QY 430 TTMRRE 435
D 431 RAFREE 436

RESULT 4
QY6N7 PRELIMINARY; PRT; 436 AA.
AC QY6N7;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE RHIZOBIOTIN SECRETION PROTEIN RSPE.
GN RSPE.
OS Rhizobium leguminosarum (biovar trifolii).
OC Plasmid PR162Y10C.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_Taxid=386;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=162Y10;
RX MEDLINE=99318631; PubMed=10388672;
RA Oresnik I.J., Twelker S., Hynes M.F.;
RT "Cloning and characterization of a Rhizobium leguminosarum gene
encoding a bacteriocin with similarities to RTX toxins.";
RL Appl. Environ. Microbiol. 65:2833-2840(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=162Y10;
RX MEDLINE=21360352; PubMed=11467725;
RA Venter A.P., Twelker S., Oresnik I.J., Hynes M.F.;

RT "Analysis of the genetic region encoding a novel rhizobiotin from
RT Rhizobium leguminosarum by. viciae strain 306.";
RL Can. J. Microbiol. 47:495-502(2001).
DR EMBL: AF141932; AAD34891.1; -.
DR InterPro: IPR002215; HLYD.
DR InterPro: IPR003997; RLXD.
DR Pfam: PF00529; HLYD; 1.
DR PRINTS: PR01490; RTXTOXIND.
KW Plasmid.
SQ SEQUENCE 436 AA; 47419 MW; 0CAE93E641BAD75E CRC64;

Query Match 27.2%; Score 589; DB 2; Length 436;
Best Local Similarity 33.9%; Pred. No. 1,7e-29;
Matches 148; Conservative 79; Mismatches 198; Indels 12; Gaps 7;

QY 4 PKIQPTDNEQAVARIGYIATLTFVGLGWAFAPLDSAVIANG--VSAEVSODVOHLE 62
D 7 PSARAIKRL-TVAALGTVLLIGVMG--GLAATRLSGAVIASGTVVDSYKPVCHOK 63
QY 63 GCMKLKILVREGGVKAGOVLELDPQANAAAGITRNQYALKAMEARLAEKDQPSI 122
D 64 GGTGQIFVKNQDIVEAGQVLIHDDTQTRANLAIVRKRLDELSARTARLAEKDGADTV 123
QY 123 SFPADITSGADPMVAAIADQAFETERTOTGOVDLMNQRLQOSEIGIDROTGL 182
D 124 EFSNDFELAKADNPQVSRISIEGKRLFADRSSRSRKQRLERIEQLKQETEGVLAQETG 183
QY 183 LKQDGLFIEDELIDRLKLYDKGLVPRPLLEARAGSLSGISGLTLADRSKAVQASDT 242
D 184 KROEIGLEKELESIORLFQDGLVPAVNAVVALOREAANLTGELSLLANEAAKGRITET 243
QY 243 QLKROIKQEFEEVSOSITETRVLAETKEVVAASDAQRIKIVSPVNGTAQNLRFET 302
D 244 ELQITQIDDDLRSEVSQDLRAESDIEGFSRLVAEDDLKRIIDVRASGGVHQLAVHA 303
QY 303 EGAVVRAAEPLVDIADEDAFVIOAHFQPTDVNHHMGVTEVRLPAFHSAGNPDERHD 362
D 304 PGAVIAPGEALMQIVPDDALLAEVKLSPTDIDQVTHLFSFASQNST--PELNG 361
QY 363 PVA--VADRIISDPQKARLFGLIVRVYKQPLHLGRVT--AGMPAQIVYGERIVQL 418
D 362 VVAGIADLTQAD--QRSGLSYVVAEVSDEMQRLQGVTPGMPPEAFIOTGERIALA 419
QY 419 YLFSPLRDLTTLTMRRE 435
D 420 YLTKEPMDQVARAFKEE 436

RESULT 5
QY679 PRELIMINARY; PRT; 439 AA.
AC QY679;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PSE (PROTEIN SECRETION PROTEIN, HLYD FAMILY).
GN PSE.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_Taxid=382;
RN [1]
RP SEQUENCE FROM N.A.
RA York G.M., Walker G.C.;
RT "The Rhizobium meliloti exo gene and prsd/prse/exsh genes encode
components of independent degradative pathways which contribute to
production of low-molecular-weight succinoglycans.";
RL Mol. Microbiol. 25:117-134(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=1021; PLASMTD=PSYMB (MEGAPLASMID 2);
RX MEDLINE=21396508; PubMed=11481431;

| | | | |
|----|---|--------------|--------------|
| | RESULT | 6 | |
| ID | Q9HYJ9 | PRELIMINARY; | PRT; 443 AA. |
| AC | Q9HYJ9; | | |
| DT | 01-MAR-2001 (TREMBLrel, 16, Created) | | |
| DT | 01-MAR-2001 (TREMBLrel, 16, last sequence update) | | |
| DT | 01-OCT-2001 (TREMBLrel, 18, last annotation update) | | |
| DE | METALLOPROTEASE SECRETION PROTEIN. | | |
| GN | HASE OR PA3405. | | |
| OS | Pseudomonas aeruginosa. | | |
| OC | Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae; | | |
| OX | Pseudomonas. | | |
| OX | NCBI_Taxid=287; | | |
| RN | {1} | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | STRAIN=ATCC 15692 / PAO1. | | |
| RX | MEDLINE=20437337; PubMed=10984043; | | |
| RA | Stover C.K., Plam X.-O.T., Erwin A.L., Mizoguchi S.D., Warren | | |
| RA | Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Legg | | |
| RA | Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yusa | | |
| RA | Brody L.L., Collier S.N., Folger K.R., Kas A., Larbig K., Lim | | |
| RA | Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T., | | |
| RA | Reizer J., Salier M.H., Hancock R.E.W., Lory S., Olson M.V.; | | |
| RT | *Complete genome sequence of Pseudomonas aeruginosa PAO1, an | | |

| | RESULT | 7 |
|--------|--|---------------------------|
| 098LG9 | 098LG9 | |
| ID | 098LG9 | PRELIMINARY; PRT; 434 AA. |
| AC | 098LG9; | |
| DT | 01-OCT-2001 (TREMBLrel. 18, Created) | |
| DT | 01-OCT-2001 (TREMBLrel. 18, Last sequence update) | |
| DT | 01-OCT-2001 (TREMBLrel. 18, Last annotation update) | |
| DE | RHIZOBIOTICIN SECRETION PROTEIN, REPE. | |
| GN | MLI1026. | |
| OS | Rhizobium loti (Mesorhizobium loti). | |
| OC | Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group; | |
| OC | Phyllobacteriaceae; Mesorhizobium. | |
| OX | NCBI_TaxID=381; | |
| RN | [1] | |
| RP | SEQUENCE FROM N.A. | |
| RC | STRAIN=MAFF303099; | |
| RX | MEDLINE=21082930; PubMed=11214968; | |
| RA | Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S., | |
| RA | Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T., | |
| RA | Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., | |
| RA | Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M., | |
| RA | Takeuchi C., Yamada M., Tabata S.; | |
| RT | Complete genome structure of the nitrogen-fixing symbiotic bacterium | |
| RT | Mesorhizobium loti." | |
| RL | DNA Res. 7:331-338(2000). | |

DR InterPro: IPR002215; HLYD.
DR InterPro: IPR000397; RCLD.
DR Pfam: PF00529; HLYD. 1
DR PRINTS: PR01490; RLYTQXIND.
DR PROSITE: PS00543; RLYD_FAMILY; 1.
KW Protease.
SEQUENCE 443 AA: 48409 MW: 7625FD7F822BB0C44 CRC64

| | | | | |
|---------------------------|-------|-------------------|------------|------------|
| Query Match | 25.08 | Score 540.5 | DB 2 | Length 443 |
| Best Local Similarity | 31.78 | Pred. No. 2.1e-26 | | |
| Matches 138; Conservative | 85; | Mismatches 178; | Indels 35; | Gaps 6; |

| | | | |
|----|-----|---|-----|
| QY | 18 | RIGIGIALTVEGLLGMNAFPLDUSAV- IANGVSAEYSODVCHLEGGMALAKILYREGK | 76 |
| Db | 25 | RIGGVLVLLGGGFLWGLLAPLDKGVPSGSVYVNAKNRAVOHPSGGSVQLOVHEGDR | 84 |
| QY | 77 | VKAGOVLELPTONAAAGITRNOYALKMEARLLAEORDRESIFFPADLTQSADPM | 136 |
| Db | 85 | VKAGOVLLIMDTVYSRQRODLRSQOGLSMAOQARLQAEORGHAIYFPALLQARREPE | 144 |
| QY | 137 | VARAIADQAOFTERRTOTIGOVDMNARLOYOSEIEIGDROTKLQOLGFIIEDELID | 196 |
| Db | 145 | VMSLTLLOOQFLTSRRALQSELALAIRESIGSOMAEIGIQSYASRORAMLOEOIGG | 204 |
| QY | 197 | LRIKIDKVLVPRPLRLAE-----ARAGSLSGSIGRPLADRASKAVOGASDPOKJROI | 249 |
| Db | 205 | MKRLAAGVYVARNNRLDLEGGYAOIDOGASDSDTNIGRL-----GROILELTKRAI | 255 |
| QY | 250 | --KOEFFBOVSQSTETFRVLRAEYETEKEVVASDAOKRIKIVPYNGTAQMLREFTEGAVV | 307 |
| Db | 256 | ORREYQKEVSSQLAEYBMRKDELDELNRKAEADLGHQYASVAGVYVGLSTVEGVI | 315 |
| QY | 308 | RAAEVLVNIABDEDAFVIGAHFOPTDVNVHMGAVTEYRLPAHFSAGNPPPERKIDPAVA | 367 |
| Db | 316 | GAGQOLMEIVSDRGLOVEARIPEVLLIDKVOVGLPEVLLFSAFMPSSTTPRVEGEVTLVGA | 375 |
| QY | 368 | DRISDPQKQARLF-----LGIVRYDVKOLPPLHNGRVATGAPPAOIVYPTGERTVLOY | 419 |
| Db | 376 | DRLTDEKSGAPYYIVRAKVSESGIQRUNGLGILRP-----GNPVGSGFIRTGSRSMNY | 427 |
| QY | 420 | LFSPRLDPLRTTMMKE | 435 |
| Db | 428 | LEKPLTDRLHIALTEE | 443 |

| RESULT | 10 | | | |
|--------|---|--------------|------|---------|
| 005198 | | | | |
| ID | 005198 | PRELIMINARY; | PRT; | 473 AA. |
| AC | 005198; | | | |
| DT | 01-JUL-1997 (TREMBLrel. 04, Created) | | | |
| DT | 01-NOV-1999 (TREMBLrel. 12, Last sequence update) | | | |
| DT | 01-JUN-2001 (TREMBLrel. 17, Last annotation update) | | | |
| DE | MEMBRANE FUSION PROTEIN. | | | |
| GN | EXP2D. | | | |
| OS | Rhizobium meliloti (Sinorhizobium meliloti). | | | |
| OC | Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group | | | |
| CC | Rhizobiaceae; Sinorhizobium. | | | |
| OX | NCBI_TaxID=382; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-EPB1; | | | |
| RX | MEDLINE=98162063; Pubmed=9501442; | | | |
| RA | Lloret J., Wulff B.B.H., Rubio J.M., Downie J.A., Bonilla I., | | | |
| RA | Rivilla R.; | | | |
| RT | "Expolysaccharide II production is regulated by salt in the | | | |
| RT | halotolerant strain Rhizobium meliloti EPB1." | | | |
| RL | Appl. Environ. Microbiol. 64:1024-1028(1998). | | | |
| DR | EMBL: Y08703; CAB41456.1; - | | | |
| DR | InterPro: IPR002215; HYD. | | | |
| DR | InterPro: IPR003997; Rtxd. | | | |
| DR | Pfam: PF00529; Hlyd. 1. | | | |
| DR | PRINTS: PR01490; RTXTOXIND. | | | |

SQ SEQUENCE 473 AA; 51603 MW; D536469E385319FF CRC64;

| | | | | |
|---------------------------|-------|----------------|----------|------------|
| Query Match | 24.8% | Score 536.5 | DB 2 | Length 473 |
| Best Local Similarity | 28.5% | Pred No. 4e-26 | | |
| Matches 119; Conservative | 96 | Mismatches 199 | Indels 3 | Gaps 2 |

| | | |
|----|-----|---|
| Qy | 20 | GXGIIATLFFVOLLMAAFAPLDSVAVINGVYSAVS--ODVUOHEGMIATLIVREGKVC 78 |
| | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : |
| Dd | 59 | GLTITLVAFGFPGCHAPSTELSSASVYSGTIIYDSKRTYSHPEGVLSNLTIVQEGDNA 118 |
| | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : |
| Qy | 79 | AGOVLEFDPTQANAAGITRNOYVALKAMEARLLAERDORPSISFPADLTISGRADPMVA 138 |
| | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : |
| Dd | 119 | PGOPLMQJEDTRARSDLOATESRRVGLIATKLARLSELAGQAVDFPDPLVA--AGEAAA 176 |
| | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : |
| Qy | 139 | RAIADQAOFTERRQTIOGOVULMNAORLOYSLEIGIDQTQGLQFIDELIDL 198 |
| | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : |
| Dd | 177 | DAVFAETAFEEKRNEAKGRITAIKRTITEEYSEKASKLTQLOLATDROIELMNEORTAIA 236 |
| | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : |
| Qy | 199 | KLYOGUVPBRLLALERAGSLSGIRGLTADBSKVAOGASDTOLKVRIOKOEFEQVS 258 |
| | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : |
| Dd | 237 | TLVEAFAPORSKRLIEIDARLSLATAIKELTAGDAQAOKAMAGAEMLTWTIESDPOSEIA 296 |
| | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : |
| Qy | 259 | QSITETRYRLAEVIEKEVVASDAOKRIKIVSPVNGTAQNLRFEEGAVVRAAEPLVDIAP 318 |
| | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : |
| Dd | 297 | GEITTARLJELAEVERIRIISAADVLRLEIRAPQGIYANTOLRPPGSAAVYPGQPLDIY 356 |
| | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : |
| Qy | 319 | EDEAFVIAHFOPITDVVDVNHGMVTEYVRLPAFHSAGNPDPERHDPPVAVADRISDPQOKAR 378 |
| | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : |
| Dd | 357 | EDEPLLVEMHVSITRDIIDISITIGSSTOIRLTAYNORSHPMDGKVTYTAADQSMDEKSNVA 416 |
| | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : |
| Qy | 379 | LFLGLVIRVDVQOLPPHLRGRTAAMPQOVIYPTGEKRVLYOLUPLSLDILTITMREE 435 |
| | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : |
| Dd | 417 | YFVARAEIAPSLANDPIDIRIPCMPEVLLVHAKARSAIDIVLSPSDSISNRAFREDE 473 |
| | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : |

```

RESULT      11
0921J2
ID      0921J2      PRELIMINARY;      PRT;      511 AA.
AC      0921J2;
DT      01-DEC-2001 (TRMBRLrel. 19, Created)
DT      01-DEC-2001 (TRMBRLrel. 19, last sequence update)
DT      01-DEC-2001 (TRMBRLrel. 19, last annotation update)
DE      ALKALINE PROTEASE SECRETION PROTEIN ABRE.
GN      APRE OR RC0428.
OS      Rickettsia conorii.
OC      Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC      Rickettsiaceae; Rickettsiae; Rickettsia.
OX      NCBI_TaxID=781;
RN      [1]
RN      SEQUENCE FROM N.A.
RC      STRAIN-MALISH 7;
RX      MEDLINE=21442074; Pubmed=1157893;
RA      Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA      Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA      Raoult D.;
RT      "Mechanisms of evolution in Rickettsia conorii and R. prowazekii."
RL      Science 293:2093-2098(2001) .
DR      EMBL; AE008606; ALN02966.1; -.
KW      Protease. Complete proteome.
SQ      SEQUENCE      511 AA;      57582 MW;      4BD9B8A55666A5C CRC64;

```

| | | | | | | |
|----|---|------------------|--------------------|------------|-------------|-----------|
| | Query Match | 24.5% | Score 530.5: | DB 16: | Length 511: | |
| | Best Local Similarity | 30.0%: | Pred. No. 1.le-25: | | | |
| | Matches 135; | Conservative 88; | Mismatches 202; | Indels 25; | Gaps 7 | |
| Oy | 5 KIORPMDNQAVAR--IGYITATFVGLG--WAFAPADSAVIANGVSAEVS-QDVQ 59 | | | | | |
| | : : : | : : : | : : : | : : : | : | : : |
| Dd | 68 KIDDKDRNNVAQAARSITLFGIYMIFVLIGLGMSALAPLDSCAIVAGIMPTNKKTIQ 127 | | | | | |
| Oy | 60 HLEGGMALILYREGGEVAGOVLFELDPDQAANAAGITRNQVALKAMEARLLAEQRDOR 119 | | | | | |
| | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |

Db 128 HHGGIINAIYVKGDKLIEETRIRIKSEHNILGQVFNFLATENRLLAERDNL 187
 QY 120 PSIFPADLTQSARADPWARAIADEQAQFTERRQTIQGOVDLMAAQLQYQSEIEGIDRQ 179
 Db 188 EQIEFSPFLMODINLPEVAKIITHQENLFRSKREYVNSEKDALHQNIAQLEKTEGLEAK 247
 QY 180 TQGLKDLQGEIEDELIDRLKLYDKGLVPRRLALAEARAGSLSSIGRLTADRKAQVGA 239
 Db 248 KVAASKTAEVAYQDRKALFLTKKEGFOVKALLDQEAQKVAASKSDVATTEAINGINHA1 307
 QY 240 SDTOLAKRQIKQEFFEVSOSITETRYRLAEVTEKEVAVASDAQRIKIVSVNGTAONLR 299
 Db 308 TETQIKINONKTERTELTELREAOVOTASLKEKYNALTDLSLRVYIIRAVDGVISLX 367
 QY 300 FEFEGAVVRAAEPLVDIAPEDEAFVIOAHQPTVDVNHMGMTVEVRLPAFHSAGNP--- 356
 Db 368 YHTIGVYISGQPIWELSPINDPLITEAKVSKQKIDSVHGLAKIRPSAKSTTPTFT 427
 QY 357 -DPERHDPVAVADRISDP-QKQARLFLGIVRV-----VKOLPRLRGVTAQMPA 405
 Db 428 GKVVISPDIVDERQYPGQOQDNVYVARVEIDMDEFNKVAKVKNLELH-----PGMQA 481
 QY 406 QVIVPTGERTVLOYLFSPRLDTLTMTREE 435
 Db 482 EVOIVTGTFTLLRYLDPVDTAFKAFREK 511

RESULT 12
 Q9XB63 ID Q9XB63 PRELIMINARY; PRT; 443 AA.
 AC Q9XB63;
 DT 01-NOV-1999 (Tremblrel. 12, Created)
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE PRTE PROTEIN.
 GN PRTE.
 OS Erwina amylovora.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 EC Erwina.
 NCBI_TaxID=552;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-E9;
 RX MEDLINE=9303693; PubMed=10373365;
 RA Zhang Y., Bak D.D., Heid H., Geider K.;
 RT "Molecular characterization of a protease secreted by Erwina
 amylovora."
 RL J. Mol. Biol. 289:1239-1251(1999).
 DR EMBL: Y19002; CAB42875.1; -
 DR InterPro: IPR002215; HLYD.
 DR InterPro: IPR000847; HTH_LYSR.
 DR InterPro: IPR003997; RtxD.
 DR Pfam: PF00529; HLYD.1.
 DR PRINTS: PR01490; RTXTOXIND.
 DR PROSITE: PS00543; HLYD_FAMILY.1.
 DR PROSITE: PS00044; HTH_LYSR_FAMILY; UNKNOWN.1.
 SQ SEQUENCE 443 AA; 48865 MW; 59AD285715994011 CRC64;

Query Match 24.1%; Score 522; DB 2; Length 443;
 Best Local Similarity 29.8%; Pred. No. 3e-25;
 Matches 137; Conservative 81; Mismatches 181; Indels 60; Gaps 9;

QY 7 QRPDNRQAVARIGYGIATLTFVGLGMAAFAPLDSAVIANGVSAE-VSDQVHLEGGM 65
 Db 15 QIPDE-RRTYRMGMGLGVGLFGLMAAFAPLDSAVSPGTVTSGNSKTVQAPASGI 73
 QY 66 LAKTIVEGEKVKAGOVLFELDPQANAAGITRNOYVALKAMEARLLAEERDQPSISFP 125
 Db 74 INNIAVEGDKVKAGALIVQLSVOQAQVDSLDRYTTTLATGRLLAERDGLNQYTF 133
 QY 126 ADLTQSARADPWARAIADEQAQFTERRQTIQGOVDLMAAQLQYQSEIEGIDRQGLKD 185

Db 134 PIFAQVQDPRVVEIIALQTLQFASRRQGLQSEID-----GYKQSMGMRFLQKGLQD 186
 QY 186 -----QLGTEIEDELIDRLKLYDKGLVPRRLALAEARAGSL-----GSIGRLTAD 231
 Db 187 SRANKQIQLSSLRQOMNSMKQLADGTLPRNRYLDVORQRAEVSSSSNENAGRIQGL--- 243
 QY 232 RSKAVQASDPTOLKVRQIKQEFFEVSOSITETRYRLAEVTEKEVAVASDAQRIKIVSPV 291
 Db 244 -QKQLQ---ESQQRIDQRFADYQREVRTQLAQTQMDISEFPNKLQMDPDLGNATISPV 299
 QY 292 NGTAQNLRFETEGAVVRAAEPLVDIAPEDEAFVIOAHQPTVDVNHMGMTVEVRLPAFH 351
 Db 300 DGTIVAGINIFTOGGVAGSHLMDVPSQALVDSHLKVDLIDKVVQGLPVDLMTAFN 359
 QY 352 SAGNDEPDERHDPVAVADRISDPQKQARLFLGIVRVQKOLPRLRGVTA----- 401
 Db 360 QNKTPKIPGTIVLSADL-----VDTSGEPPYRMQVTVSPGMAMLSG 404
 QY 402 -----GMPAQIVPTGERTVLOYLFSPRLDTLTMTREE 435
 Db 405 EDIKGMPVEVFAVTKGSRSLSTYLFKPIIDRAHNTLREE 443

RESULT 13
 Q54457 ID Q54457 PRELIMINARY; PRT; 443 AA.
 AC Q54457;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE LIPI.
 GN LIPI.
 OS Serratia marcescens.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Serratia.
 NCBI_TaxID=615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SR41;
 RX MEDLINE=96062219; PubMed=7592412.
 RA Akatsuka H., Kawai E., Omori K., Shibatani T.;
 RT "The three genes lipB, lipC and lipD, involved in the extracellular
 secretion of the serratia marcescens lipase which lacks an N-terminal
 signal peptide."
 RL J. Bacteriol. 177:6381-6389(1995).
 DR EMBL: D49826; BA08632.1; -
 DR InterPro: IPR002215; HLYD.
 DR InterPro: IPR003997; RtxD.
 DR Pfam: PF00529; HLYD.1.
 DR PRINTS: PR01490; RTXTOXIND.
 DR PROSITE: PS00543; HLYD_FAMILY.1.
 SQ SEQUENCE 443 AA; 49000 MW; B69DEC70F474EE46 CRC64;

Query Match 24.1%; Score 521.5; DB 2; Length 443;
 Best Local Similarity 29.3%; Pred. No. 3.e-25;
 Matches 133; Conservative 88; Mismatches 194; Indels 39; Gaps 7;

QY 6 IQRPDTNF-----QAVARIGYGIATLTFVGLGMAAFAPLDSAVIANGVSAEVS-Q 56
 Db 5 IGEPQDSYTEIEIPQDERFTRMGMVVGIGLFLMAAFAPLDSAVSPGTVTSGNRK 64
 QY 57 DVQHLLEGMLAKIIVREGKVKAGOVLFELDPQANAAGITRNOYVALKAMEARLLAEER 116
 Db 65 TVQAPASGIININIAVRQGDVKAGEVILVQLSVOQAQVDSLDRQYTTTLATGRLLAER 124
 QY 117 DQRPISPADLTQSARADPWARAIADEQAQFTERRQTIQGOVDLMAAQLQYQSEIEGI 176
 Db 125 DGLSTIVTFSPILDVAKKPRVAEIIALQTLQFASRRQGLQSEID-----GYQSMGDI 177
 QY 177 DRQTOGLKD-----QLGTEIEDELIDRLKLYDKGLVPRRLALAEARAGSLSSIGRLT 229
 Db 178 RFLQKGLQDSGRNQIQLSSLRQOMNSMKQLAAGDGLPRNRYLEVORQFAEVNSSIDETV 237

```
QY 230 ADRSAVQASDTOLKVRQIKOEFEVQSITETRVRLAEVTEKEVVASDAQKRIKIVS 289
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 238 GRIGQLOKOLLESQORIDQREAVQREVRLQATQMDASEFRKLMADPDLGNATITS 297
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 290 PVNGTAONLRFTEGAVVRAEPLVDIAPDEDAFVIOAHFOPTDVNVHMGMTVEVRLPA 349
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 298 PVDGTAVGLNFTQGVGAGDHLMDVPPSQATLVDSRLKVDLEDRVNGLPVLDLMTA 357
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 350 FHSAGNDPERHDPVAVADRISD-----PQKQARFL--GIVRVADKOLPRLRGVTA 401
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 358 FNQKTKPICTVTLVSDRLVDKANGEPYQMOGYTVSPESMKMLSGEDIKP----- 409
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 402 GMPAQVIVPTGERTVLOYLESPRLDRLRTMREE 435
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 410 GMPVEVFKTGSRLSLSYLFPKPLDRAHTSLTEE 443
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
Q9ZDL4 PRELIMINARY; PRT; 511 AA.
ID 09ZDL4;
AC 09ZDL4;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, last sequence update)
DE 01-DEC-2001 (TReMBLrel. 19, last annotation update)
DE ALKALINE PROTEASE SECRETION PROTEIN APRE (APRE).
GN RP314.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MADRID E.
RX MEDLINE=99039499; Pubmed=9823893;
RA Andersson S.G.E., Zomrodipour A., Andersson J.O., Naeslund A.K.,
RA Slicheritz-Ponten T., Alsmark U.C.M., Podowski R.M.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria."
RL Nature 396:133-140(1998).
DR EMBL; AJ235271; CAAL4774.1; -
DR HSSP; P10802; 11YU.
DR InterPro; IPR002215; HLYD.
DR InterPro; IPR003997; REXD.
DR Pfam; PF00529; HLYD; 1.
DR PRINTS; PR01490; RHTOXIND.
DR PROSITE; PS00543; HLYD_FAMILY; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 511 AA; 57770 MW; CFC6EC6005BA34A8 CRC64;

Query Match 24.0%; Score 519.5; DB 16; Length 511;
Best Local Similarity 29.3%; Pred. No. 5.3e-25;
Matches 134; Conservative 93; Mismatches 192; Indels 39; Gaps 10;
```

```
Db 304 RHATTEQIKIINHONKTYTERTLTLEADQITASLKEKNSLTLNRIYSPVDGIY 363
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 296 ONLFEFEGAVVRAEPLVDIAPDEDAFVIOAHFOPTDVNVHMGMTVEVRLPAFHSAGN 355
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 364 NNLKHTIGVISHGQIPIMELSPINDPLIETARISQKNIDSVHEGLVAKIRFSAFSRT 423
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 356 P-----DPERHDPVAVADRISDPQKQARFLGIVRD-----VKOLPRLRG 397
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 424 PTFGKVVATSPD----IVDEROHLGQODNYVARIETIMEEFNKVKARKMLAH--- 476
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 398 RVTGMPAQVIVPTGERTVLOYLESPRLDRLRTMREE 435
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 477 ---PGKQAEVQIVGTTRTLRLTLDPVTDPAFAFREK 511
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15.
ID 005694 PRELIMINARY; PRT; 435 AA.
AC 005694;
DT 01-JUL-1997 (TReMBLrel. 04, Created)
DT 01-JUL-1997 (TReMBLrel. 04, last sequence update)
DE 01-DEC-2001 (TReMBLrel. 19, last annotation update)
DE PSE PROTEIN.
GN PSE.
OS Rhizobium leguminosarum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=384;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8401;
RA Finnie C., Hartley N.M., Findlay K.C., Downie J.A.;
RT "The Rhizobium leguminosarum prsD genes are required for secretion of
RT several proteins, some of which influence nodulation, symbiotic
RT nitrogen fixation and exopolysaccharide modification."
RL Mol. Microbiol. 0:0-0(0).
DR EMBL; Y12758; CAJ73293.1; -
DR InterPro; IPR002215; HLYD.
DR InterPro; IPR003997; REXD.
DR Pfam; PF00529; HLYD; 1.
DR PRINTS; PR01490; RHTOXIND.
DR PROSITE; PR01490; RHTOXIND.
SQ SEQUENCE 435 AA; 47533 MW; DA1F71655D96D131 CRC64;

Query Match 23.6%; Score 512; DB 2; Length 435;
Best Local Similarity 30.8%; Pred. No. 1.3e-24;
Matches 132; Conservative 71; Mismatches 208; Indels 18; Gaps 4;
```

Mon Jul 22 09:14:02 2002

us-09-913-414-5.rsp

Page 9

[illegible]

Search completed: July 19, 2002, 10:48:43
Job time: 418 sec

THIS PAGE BLANK (USPTO)